6

Journal of Mathematics and Physics

(Founded by C. L. E. Moore)

Volume XIV

1935

Editors

VANNEVAR BUSH
FREDERICK G. KEYES
DIRK J. STRUIK
MANUEL S. VALLARTA
NORBERT WIENER
FREDERICK S. WOODS

Managing Editor

PHILIP FRANKLIN

Published by the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY PRESS

QA1
J87
v.14

OHIO STATE
UNIVERSITY

CONTENTS

Five Papers on the Theory of Probability. *By D. J. Struik*..... 1
Remarks on Causality and Probability. *By E. Hopf*..... 4
The Probability of Position in a Canonical Ensemble. *By G. E. Uhlenbeck* 10
Random Functions. *By N. Wiener*..... 17
Some Analytical Problems Relating to Probability. *By B. Jessen*..... 24
Principles of Probability in Natural Science. *By Felix Bernstein*..... 28
Sur les Moyennes Arithmétiques des Séries-Noyaux des Développements en Séries D'Hermite et de Laguerre et sur Celles de ces Séries-Noyaux Derivées Terme a Terme. *Par Ervand Kogbetliantz*..... 37
The Lift and the Moment of an Arbitrary Aerofoil-Joukovsky Potential. *By J. G. Estes*..... 100
Some Problems of Mathematical Biology. *By J. B. S. Haldane*..... 125
Applications of the Fourier Transform Theorem. *By Albert E. Heins*..... 137
On S, P, D and F Biharmonic Polynomials. *By Frank L. Hitchcock*..... 143
On Minkowski's Definition of Length and Area. *By Philip Franklin*..... 179
On the Closure of the Set $\{e^{in\theta}\}$. *By Morris L. Kales*..... 187
A Simple Non-linear Integral Equation. *By Morris L. Kales*..... 189
The Range of de la Vallee Poussin Summation. *By G. Rutledge and R. D. Douglass*..... 191
Fourier Transform Solution of Equations. *By Norman Levinson*..... 195
Note of Heawood Color Formula. *By I. N. Kagno*..... 228
Applications of Tauberian Theorems to the Study of Entire Functions. *By Bernard Friedman*..... 232
Reaction Velocity in the System: $\text{Solid}_{\text{I}} + \text{Gas} \rightarrow \text{Solid}_{\text{II}}$ *By George S. Gardner*..... 241
Note on the Absorption Spectra of Oxygen at High Concentration. *By B. J. Eiseman, Jr.*..... 260
A Particular Boundary Value Problem. *By Maurice J. Fish*..... 262
Geodesics on Polyhedral Surfaces. *By P. Franklin*..... 274
The Geometry of Linear Homogeneous Partial Differential Equations of the First Order. *By G. P. Wadsworth*..... 279
The Equivalent Wave Method—The Use of Vectors in Studying Electrical Transients. *By F. Creedy*..... 291
Matrices in Electric Circuit Theory. *By Richard Stevens Burington*..... 325
Almost Periodic Functions and the Vibrating Membrane. *By Arra Steve Avakian*..... 350

514471

FIVE PAPERS ON THE THEORY OF PROBABILITY

By D. J. Struik

The five papers which follow were read at the December 1933 meeting of the American Mathematical Society at Cambridge, Massachusetts, which devoted one morning to the discussion of the theory of probability. The papers represent some of the varied topics in which the theory of probability plays a rôle. It would have been easy to increase the number of subjects by including papers on life insurance, several branches of engineering or the classical subjects relating to gambling. But already the range of topics is imposingly wide, embracing such seemingly remote subjects as Riemann's theory of the ζ-function, the theory of the atom and the theory of chromozomes. The problem arises as to the common principles underlying these different branches of science and the mathematical theory which enables us to express these common principles.

There is here a certain analogy with what happened in the seventeenth century. A great number of simple machines had been invented, a great deal of engineering had been done, a great deal of research on the mechanics of rigid and even of fluid bodies was the result. Tartaglia had studied the motion of projectiles, Galileo the motion of freely falling bodies, Stevin the motion on the inclined plane, Kepler the motion of the heavenly bodies. During the seventeenth century this process continued in the natural sciences and in egineering. The contemporary research in mathematics followed the requirements of the age and developed methods dealing with the analysis of motion in its different aspects, especially in the form of the tangent problem of curves. The methods were first all ad hoc, without much of a general principle, till Leibniz at last discovered not only the general method, but also the mathematical symbolism needed to express the relations. That was the invention of the calculus, which created an adequate algorithm to cope with all such phenomena in which small changes play the essential rôle.

At present we have not the difficulty of coping with regularities in phenomena subjected to small changes. Probability phenomena all deal with regularity in events which are repeated a great number of times. Classical mechanics in its different forms and classical physics dealt with changes of a local character. In probability phenomena we

deal with changes in the long run, but only such changes as show a certain (statistical) regularity. The problem is to define the common character more accurately and to find the mathematical algorithm which expresses this particular type of regularity.

The necessary mathematical theories became available with the advance of the theory of point sets and the measure theory of Lebesgue. This allowed Borel to introduce notions of measure into the theory of probability.[1] The notion "almost always," which is so essential for probability phenomena, could be connected with the very definite mathematical conception "always but for a set of measure zero."

This opened a field of research which introduced probability notions into the field of pure analysis, and gave them an exact meaning. The axiomatics of probability now became a chapter in the theory of point sets, as is witnessed by Kolmogoroff's recent book.[2] The paper by Dr. B. Jessen, read at the meeting, gives an account of some modern aspects of this type of work. The opening lines of his paper, which state that the words "probability and expectation" may be just as well replaced by the words "measure and integration" give an insight into the nature of this formal side of the theory of probabilities. We can, in the same way, use in calculus words as "velocity," "rate of change," etc. instead of the formal word "derivative."

In the application to problems of natural science two ways are open. The first is a purely descriptive one, in which we try to follow the natural phenomena with an adequate mathematical apparatus without attempting an explanation in causal terms. Dr. Wiener, who developed such an apparatus in the case of the Brownian motion in one of the first theories which showed the necessity of the introduction of the notion of measure into the theory of probability, gives in his paper a new account of his theory.

The other way is to explain probability phenomena by connecting them with the laws which govern the motion of the participating bodies. The main transition from mathematics to natural science lies in Leibniz' principle of sufficient reasoning. Is there a "sufficient reason" in statistics? The simplest case seems to be that of the mechanics of rigid bodies. Since the time of Maxwell and Boltzmann this problem has been known as the ergodic theory, but only in the last years has considerable advance been made. It seems that Caratheodory[3] for the

[1] E. Borel, Rend. Circ. Mat. Palermo 27 (1909), p. 247–271.

[2] A. Kolmogoroff, Ergebnisse der Mathematik II, 3, 1933.

[3] C. Caratheodory, Sitzungsber. preuss. Akad. d. Wiss., 1919, p. 580; see G. D. Birkhoff and B. O. Koopman, Proc. Nat. Acad. Sc. U. S. A. 18, 1932, p. 279.

first time introduced measure considerations into rigid dynamics. At the same time von Smoluchowski,[4] in a paper in the "Plank Festschrift," expressed the general principle that probability phenomena must be of a causal character, at the same time expressing the conviction that the type of relations which lead to probability phenomena are those for which small initial causes lead to large results. v. Smoluchowski could not carry out his idea, not having the mathematical apparatus. Dr. Hopf shows in his paper how Caratheodory's and v. Smoluchowski's ideas combined lead to a satisfactory explanation of the classical probability phenomena. Dr. Bernstein shows us additional insight into the origin of v. Smoluchowski's ideas. The causal theory behind Wiener's theory is Einstein's theory of the Brownian motion.

Is it possible to introduce similar principles into other fields where the theory of probabilities finds an application? The paper of Dr. Bernstein gives hope that in biology this task can be accomplished. Not all properties of the dynamical laws are used in the proof of the ergodic theorems of dynamics. In the same way, we do not need to understand the complete way of formation of the masses of biological entities. The clear-cut way in which in several cases the conditions of a Laplace urn are satisfied gives the hope that a dynamical basis for the foundations of probability in biology can also be found.

Dr. Uhlenbeck's paper, at last, gives valuable suggestions as to the connection between the statistical phenomena of quantum mechanics and those of classical mechanics. Here a relation of quantum dynamics and quantum statistics exists through the relation between the Schrödinger equation and the probabilities derived from them. A greater difficulty is offered by the relation of quantum theory and classical theory. The existence of such relations, as goes forth from Dr. Uhlenbeck's paper, carries causality relations also into quantum relations and helps us to understand probability phenomena as belonging to one single group of events characterised in a similar way. In this way we may carry the hope that the theory of probability will soon have left the state in which mechanics was before Galileo.

[4] M. v. Smoluchowski, Plank Festschrift, Naturwissenschaften 6 (1918), p. 253.

REMARKS ON CAUSALITY AND PROBABILITY

By E. Hopf

Most mathematicians have grown accustomed to regard the calculus or probabilities as a branch of pure mathematics, in particular, of the theory of measure. Yet it cannot be denied that it belongs to natural science as well since it deals with certain mass phenomena. In order to give the theory of probabilities a satisfactory background it is necessary to penetrate the nature of those phenomena. The classical theory resting upon the notion of "equally likely" has achieved little in this direction, as d'Alembert's paradox and the cases involving continuous probabilities show. As long as our interest is directed toward the practical application of the result we might be satisfied by the success of our calculation. The question, however, why in nature certain cases are produced with equal likelihood remains unanswered. It should be mentioned that the same holds of the modern frequency theory of probabilities.

The fact that in a long series of throws of a perfect coin head and tail occur approximately equally often is usually attributed to the symmetry. Suppose that in such a series head turns up with the relative frequency λ. Does it follow from the symmetry that $\lambda = \frac{1}{2}$? No, all we can infer is that there exists another equally possible series (head and tail interchanged) in which the relative frequency of head equals $1 - \lambda$. The really important fact about the actual experiment is thus, that head always turns up with the same frequency no matter how the coin is tossed. We could have generally convinced ourselves of this if we had taken an asymmetric die. It is a characteristic feature of all frequency phenomena underlying the probability theory that the observed frequencies are practically constant under varying circumstances. Forecasting would be impossible if this were otherwise. The importance of that regularity property, for a deeper insight into the origin of the probability laws, has recently been emphasized by my colleague D. J. Struik.

The causes producing the events in question must be of a particular type in order to make the mentioned regularity possible. To study this type of causal relationship we choose the example of the coin. Denote

by P the initial phase (position and velocity) of the coin when released. The coin bounces on the floor and finally comes to rest turning up head or tail. The phase space Ω is accordingly divided into two parts, H and T, leading to the event "head up" or "tail up" respectively. It is not unreasonable to assume strict causal relationship between initial phase and final position, and to suppose this relationship to be independent of the time. H and T are thus well defined, distinct parts of $\Omega = H + T$. An infinite sequence of throws represents an infinite succession of initial phases, $P_1, P_2, P_3, \ldots$. Let A be a simple but arbitrary part of Ω (sphere or parallelepided). There will be a definite fraction of the P_i falling into A, i.e., we may assume that the limit of

$$\frac{1}{n}\sum_{1}^{n}\varphi_A(P_i)$$

exists as $n \to \infty$, $\varphi_A(P)$ being the characteristic function of the point set A. The limit

$$\mu(A)$$

will be an additive point set function of A. Suppose furthermore that

$$\mu(A) = \int_A f(P)dm\,,$$

m being a given volume measure in Ω in the sense of Lebesgue, and $f(P)$ being a non-negative function. $f(P)dm$ is, so to say, the frequency with which the P_i fall into dm. It is evident that $f(P)$ is (within wide limits) an entirely arbitrary function, since the sequence P_i is arbitrary to a large degree. Experience suggests only that $f(P)$ satisfies the following loose restriction. In a small region A, the points P_i will in general be uniformly distributed. This means that in general $f(P)$ varies little within such a region. There is, however, no further restriction upon $f(P)$.

The relative frequency with which the points P_i fall into the region H, i.e., with which the coin turns up head, equals

$$\mu(H) = \int_H f(P)dm\,; \qquad \int_\Omega f(P)dm = 1\,.$$

How is it possible that this expression varies little when $f(P)$ is "practically" arbitrary (in the sense outlined above)? Let us choose $f(P)$ equal to zero outside a region A and constant inside. This is practically legitimate if A is not too small. The latter condition

mirrors itself in the fact that we cannot systematically confine the initial phase of the coin to too small a region. We obtain then

$$\mu(H) = \frac{m(AH)}{m(A)} .$$

In order that the relative frequency of the event "head up" be practically independent of the particular manner of throwing the coin, it is thus necessary that the part H of Ω be practically uniformly distributed in Ω, i.e. that H shares an approximately constant fraction (in measure) with every not too small region A. Conversely, uniform distribution of H implies in turn constancy of the relative frequency of the event H. This loose type of reasoning is inevitable, as we are dealing with but approximate phenomena. It indicates, nevertheless, clearly enough the type of causal relationship responsible for the phenomenon of statistical regularity. The preceding argument is not entirely novel, Poincaré[1] and Smoluchowski[2] have fully recognized its bearing upon the foundations of probability theory. To Poincaré we owe the beautiful explanation of the roulette game: The wheel is divided into a large number of alternatively red and black sectors of equal length. The frequency with which the wheel stops at a red sector is always approximately one half because the totality of these sectors is nearly uniformly distributed on the wheel (with density one half). Other examples have been given by Hostinski.

It is worth while to trace the consequences of the above argument somewhat further. So far we have realized that the statistical regularity of an event is equivalent to the practically uniform distribution of a certain point set connected with this event. Suppose two persons toss each a coin independently and repeat this experiment infinitely often. What will be the relative frequency of the simultaneous event HH or of the other combinations, HT, TH, TT? We now have a sequence of pairs of initial phases, $P_1, Q_1; P_2, Q_2; \ldots$. First of all, we must explain what we mean by independence. The crude interpretation is that there is no functional relation between P and Q. We suppose, similarly, that the pairs P_i, Q_i occur with a certain frequency in the elementary regions $A \times B$ of the product space $\Omega \times \Omega$,

$$\mu(A \times B) = \lim \frac{1}{n} \sum_{1}^{n} \varphi_A(P_i)\, \varphi_B(Q_i) .$$

[1] H. Poincaré, Calcul des Probabilité's, Paris, 1912, Introduction.
[2] M. v. Smoluchowski, Naturwissenschaften 6 (1918), p. 253.

The precise meaning of the independence is that $\mu(A \times B)$ can be written in the form

$$\mu(A \times B) = \iint_{A \times B} f(P, Q) dm_P \, dm_Q ,$$

f being summable over $\Omega \times \Omega$ (the (Pi, Qi) are really "spread out" on $\Omega \times \Omega$). It must be emphasized that $f(P, Q)$ is arbitrary to a large extent and that in general

$$\mu(A \times B) \neq \mu(A) \, \mu(B) .$$

From the approximately uniform distribution of H in Ω, however, we may infer that the set $H \times H$ corresponding to the compound event $H \times H$ is nearly uniformly distributed in $\Omega \times \Omega$. Hence

$$\mu(H \times H) = \mu(H) \, \mu(H)$$

will approximately hold independent of the distribution of the Pi, Qi, i.e. of the special manner of throwing. Similar arguments apply to the other combinations HT etc., and to the combinations of more than two events. The law of compound probabilities thus appears as a consequence of the notion of statistic regularity. Instead of starting with sequences P_i and P_i, Q_i, we might begin with continuous distributions of $f(P)$ and $f(P, Q)$ and show afterward that the events in question occur statistically regularly within "most sequences" ("law of large numbers"). It seems, however, desirable to choose first the former way because the details involved are brought out more clearly.

Looking back, we state once more the criterion for statistic regularity of an event produced by a certain causal mechanism. The part of the phase space Ω corresponding to the event in question must be practically uniformly distributed in Ω. Moreover, this uniformity takes already place within regions in which the distribution functions $f(P)$ vary little. Two things are thus responsible for statistic regularity. A particular type of causal mechanism governing the phenomenon in question, and sufficient continuity of the initial distribution functions involved. At present only few examples of such mechanisms are accessible to mathematical treatment, but there is no doubt that nature provides a large variety of mixing processes producing statistic regularity. This is also the reason why we can consider our distribution functions sufficiently continuous. We can do nothing better than quote Poincaré's own words for this (Science et Méthode): "In the course of history complex causes have worked a great while; they have contributed to produce the mixture of elements and they have tended to make every-

thing uniform at least within a small region; they have rounded off the corners, smoothed down the hills and filled up the valleys. However capricious and irregular may have been the primitive curve given over to them, they have worked so much toward making it regular that finally they deliver over to us a continuous curve. And this is why we may in all confidence assume its continuity."

So far our considerations have been rather loose, owing to the but approximate nature of our subject. If we attempt to render the theory exact we encounter a slight difficulty in the fact that, apart from trivial exceptions, perfectly uniformly distributed point sets do not exist. The theory must therefore be based upon limit relations. There are parameters in the problems that, when tending toward certain limits, make the event sets considered more and more homogeneously distributed. In the actual case these parameters have, of course, constant values differing sufficiently little from their respective limits. In Poincaré's case of the roulette the parameter is the number n of sectors $n \to \infty$. In the case of the asteroids, it is the time, $t \to \infty$. A few more words should be said about the problems of the coin and of the die. There are two different possibilities of interpretation. Either we consider the energy range fixed and the numbers μ and $1 - e$ small, μ and e being the coefficients of friction and elasticity respectively. This means that the bouncing of the coin and die on the floor is considered nearly conservative. We might, on the other hand, imagine that μ and e being kept fixed, the largeness of the energy with which the coin is released is responsible for the regularity phenomenon. The energy depends upon the initial phase, but it can be made to appear as a parameter after a suitable transformation. There is thus a certain freedom in the construction of mathematical models, and it is not easy to decide which one is closest to the actual case. At any rate, the common scheme brought out might serve for an exact theory of statistic regularity.

In all cases we have a certain space Ω of initial phases P. The phases P which lead to a certain event E constitute a certain part of Ω. The phases P might be tied up with the final event directly (as in Poincaré's remark about the roulette), or by the mediation of a causal mechanism (coin, die, asteroids, etc.). The part E_α of Ω which corresponds to the event E will depend upon one or several parameters α. Let m denote a finite volume measure in Ω. We say then that the event E is statistically regular under the circumstances considered if, for any part A of Ω,

$$L(E) = \lim_{\alpha=\alpha_0} \frac{m(AE_\alpha)}{m(A)}$$

exists and if $L(E)$ is independent of A. It follows then that, for any function $f(P)$ summable over Ω,

$$\lim_{a \to a_0} \frac{\int_{E_a} f(P)\, dm}{\int_{\Omega} f(P)\, dm} = L(E)\,.$$

Using modern terminology we can formulate this briefly by saying that the characteristic functions of the event sets E_a converge "weakly" toward a constant limit $L(E)$. $L(E)$ is called the relative frequency of the event E. It is easily verified that $L(E)$ is independent of the special measure m.

The theory resting upon this notion is developed in more detail in the author's paper[3] on probability, statistics and causality. Chief attention has there been devoted to statistical regularity produced by conservative mechanisms (the parameter being the time t). The main purpose of this was to tie the theory up with the ergodic theory. It is, however, evident that the theorems on compound probabilities and on the frequency in sequences remain true when based upon the above general notion.

[3] E. Hopf, this Journal 13 (1934), p. 51.

THE PROBABILITY OF POSITION IN A CANONICAL ENSEMBLE

By G. E. Uhlenbeck[1]

1. *Introduction.* The recent mathematical investigations of the ergodic theorem and of the foundations of the probability calculus promise to make completely clear the relation between the classical mechanics and the classical statistical mechanics of Boltzmann and Gibbs. For the physicist this relation was in general quite clear, especially since the work of Ehrenfest and v. Smoluchowski, although he realized that in statistical mechanics certain probability assumptions had to be made in addition to the classical equations of motions. It is very satisfactory that, as it seems now, one is able to *prove* these assumptions on the basis of the classical mechanics. They now become *theorems*, which will hold strictly in the limit as the time goes to infinite. I have the impression that the situation with regard to the relation between the quantum mechanics and the quantum statistics is quite analogous. The fundamental assumptions of the latter (say the "Ansatz" of the canonical ensemble, the equality of the weights of not degenerated quantum states) seem to become limit theorems (for $t \to \infty$) which follow from the equation of motion, that is the Schrödinger wave equation.[2]

On the other hand there is the problem of the relation between the quantum theory and the classical theory itself. In the development of the quantum theory the study of this relation has been a very fruitful way of approach. One has only to think of the correspondence principle of Bohr. Since the consolidation of the quantum mechanics we know that the quantum theory can be considered as a natural generalization of the classical theory. The latter is contained in the first as a limiting case if the Planck constant h goes to zero. The more precise investigation, especially of the way how this limit is reached, gives rise to, I think, quite interesting mathematical problems, which in several cases have not yet been worked out in all details.[3]

[1] From the University of Michigan.

[2] See especially J. v. Neumann, Mathematische Grundlage der Quantenmechanik, Ch. V.

[3] Perhaps I may call attention here to a few of these problems:

1. The study of the so called W. K. B. approximation method of the wave equa-

Quite analogously is the classical statistical mechanics a limiting case of the quantum statistics. The only difference is that the first will go over in the latter not only if $h \to 0$, but also if the temperature T goes to infinity. Again the question arises how this limit is reached. I will consider in the following a special problem of this kind, which has interested me in the last year or so, and which has some perhaps interesting mathematical sidelights.

2. *Formulation of the problem.* For a masspoint m in an outside force-field $V(xyz)$ the Schrödinger equation is:

$$H\psi = E\psi \qquad H = -\frac{h^2}{8\pi^2 m}\Delta + V \tag{1}$$

Let the eigenvalues be E_n (the spectrum may be discrete or continuous or both), the normalized eigenfunctions $\psi_n(xyz)$. Then $|\psi_n|^2\, dx\, dy\, dz$ gives the probability to find the particle in $dx\, dy\, dz$ if it is in the state E_n. For a canonical ensemble of these masspoints, the unnormalized probability density of position will then be:

$$S_{qu}(xyz) = \sum_n e^{-E_n/kT} |\psi_n|^2 \tag{2}$$

For a continuous spectrum this will become of course an integral instead of a sum. I am used to calling (2) the Slater sum, because Slater made some stimulating remarks about it a few years ago.[4] Of course it was already considered earlier, especially by von Neumann.

---

tion (see e.g. Pauli, Handb. der Physik, 2nd Ed. Vol. 24, p. 166). This method, the principle of which was already well known in mathematics (Jeffreys, Proc. Lond. Math. Soc., Ser. 2, **23**, Part 6), has been developed especially by Kramers, so that for *separable* mechanical systems the transition from the Schrödinger equation through the old quantum theory of Bohr to the classical Hamilton-Jacobi equation is now clear, although mathematically not yet quite sharp. For *non-separable* systems the transition has not been worked out so completely. The problem is a special case of the following question:

2. What is the relation between the integration theories of the classical mechanics and the methods of solution of the wave equation? For instance, if a mechanical problem is soluble by quadratures, can then the corresponding wave equation be reduced to ordinary differential equations of the second order? And is the analytic nature of the integrals connected with the analytic nature of the differential equations?

3. More special problems appear in the quantum theory of collisions. The series which one obtains for the scattering say of electrons by a force field have to go over for $h \to 0$ into the classical expressions. (See e.g. Mott and Massey, The Theory of Atomic Collisions, esp. Ch. VII, §3.) How the limit is reached is also here not quite clear.

[4] J. C. Slater, Phys. Rev. **38**, 237, 1931.

Now in the classical Boltzmann-Gibbs statistical mechanics we have for the probability density in phase space:

$$e^{-H(p,\,x)/kT}\,dx\,dy\,dz\,dp_x\,dp_y\,dp_z \tag{3}$$

where $H(p, x)$ is the Hamilton function. This has no direct analogy in the quantum mechanics because of the uncertainty principle. But in our case, as is usually, $H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V$ and we can integrate over the impulses to obtain the classical probability density of position. To be consistent we must take in quantum and classical theory the a priori weights analogously, which introduces a factor $1/h^3$ in (3). So we get:

$$S_{\text{class.}}(xyz) = \frac{1}{h^3}\int_{-\infty}^{+\infty}\!\!\int\!\!\int dp_x\,dp_y\,dp_z\,e^{-H/kT} = \left(\frac{2\pi mkT}{h^2}\right)^{3/2} e^{-V/kT} \tag{4}$$

Now I can enunciate the problem: *How does* (2) *contain* (4) *as a limit and how is the limit reached.*

As mentioned in §1 one knows that (2) will go over into (4) for $h \to 0$ and for $T \to \infty$. How it will go will depend very strongly on how fast V varies. From physical considerations one may further expect the following. The h and T will occur in the combination $\lambda = h/(mkT)^{\frac{1}{2}}$, which is the de Broglie wave length corresponding to a mean temperature motion. Suppose V is due to an obstacle at the origin of certain "dimension" d, then:

$$S_{qu} = S_{\text{class.}}\left(1 + c_1\frac{\lambda}{d} + c_2\frac{\lambda^2}{d^2} + \cdots\right) \tag{5}$$

which will probably converge well only for $\lambda/d < 1$. Of course d has no fixed meaning; in general instead of $1/d$ we can expect at a certain point differential quotients of V with respect to the coordinates. The bigger these are, measured always with λ as unit of length and kT as unit of energy, the slower the approach to $S_{\text{class.}}$ will be. Still an "obstacle" with certain dimensions (a sphere for instance) is nearly what will be the case if V is the potential between two gas molecules; (1) will then describe their *relative* motion and S is the probability to find them a certain distance apart. For the lighter gases, like *He*, λ/d is even at ordinary temperatures not very small. It is one at $T = 75°$. For these gases therefore the results of the classical kinetic theory will have to be revised considerably.

3. *Special cases.* Before making (5) more precise, I might mention that for special cases (2) gives quite interesting series, which sometimes can be summed exactly. This is for instance the case for the one dimensional motion and $V = -\omega x^2$ (harmonic oscillator). With some changes in variables one gets:

$$\sum_{n=0}^{\infty} e^{-n\vartheta} \frac{H_n^2(y)}{2^n\, n!} = (1 - e^{-2\vartheta})^{-\frac{1}{2}}\, e^{-y^2(\operatorname{tgh} \frac{\vartheta}{2} - 1)} \tag{6}$$

where $H_n(y)$ is the n^{th} Hermite polynomial. In the three dimensional case, if V is a monotonous decreasing function of the distance r, (2) can be reduced to:

$$S_{cu}(\tau) = \frac{1}{\pi^2 \tau^2} \int_0^{\infty} dk\, e^{-\lambda^2 k^2} \sum_{l=0}^{\infty} (l + \tfrac{1}{2})\, v_l^2\, (k, \tau) \tag{7}$$

where $v_l\,(k, \tau)$ is that solution of the radial Schrödinger equation

$$\frac{d^2\psi}{d\tau^2} + \left\{k^2 - \frac{8\pi^2 m}{h^2}\, V(\tau) - \frac{l(l+1)}{\tau^2}\right\} \psi = 0$$

which fulfills the boundary condition at the origin, and which for large behaves as $\cos\,(k\tau - \delta_l)$. For special cases, e.g. if V is the potential of an elastic sphere of radius σ at the origin ($V(\tau) = 0$ for $\tau > \sigma$; $V = \infty$ for $\tau \leqq \sigma$) one can find $v_l(k, \tau)$ and discuss (7) further.

Further if V has certain symmetry properties the eigenfunctions ψ_n will have corresponding ones. So e.g. in the one dimensional case if $V(x)$ is even, $\psi_n(x)$ will be alternatively even and odd. One can then ask what the sum (2) will be if we take *only* those terms for which the ψ_n have a certain symmetry property e.g. only those where $\psi_n(x)$ is even. In special cases one can again sum the series exactly; so one gets analogous to (6).

$$\sum_{n\ \text{even}} e^{-n\vartheta} \frac{H_n^2(y)}{2^n\, n!} = \tfrac{1}{2}\,(1 - e^{-2\vartheta})^{-\frac{1}{2}}\, e^{-y^2\,(\operatorname{tgh} \frac{\vartheta}{2} - 1)} \left(1 + e^{-\frac{2y^2}{\sin h\,\vartheta}}\right)$$

In the case of a central field $V(\tau)$ the choice of those terms in (2) for which $\psi_n(xyz) = \psi_n(-x, -y, -z)$, reduces to summing in (7) over the even values of l only. These are the changes in the general problem if instead of the Boltzmann-Gibbs statistics the Einstein-Bose or the Fermi-Dirac statistics is the valid one.

4. *The general case.* The more precise development[5] starts as follows. Write (2):

$$S_{qu} = \sum_n \psi_n e^{-H/kT} \psi_n$$

where H is the Schrödinger operator (1). Now by the theorem of the invariance of the trace or spur we may replace the (unknown!) ψ_n by any complete set of orthogonal functions. For instance we can take the wave functions of the completely *free* particle or in other words the plane waves:

$$\phi_{p_x p_y p_z} = \frac{1}{\sqrt{\Omega}} e^{\frac{2\pi i}{h}(p_x x + p_y y + p_z z)}$$

which are normalized in a big volume Ω. Taking $\Omega \to 0$, the summation over n becoming an integration over the $p_x\ p_y\ p_z$ we get:

$$S_{qu} = \frac{1}{h^3} \int\int\int_{-\infty}^{+\infty} dp_x\, dp_y\, dp_z\, e^{-\frac{2\pi i}{h}(p\cdot r)}\, e^{-H/kT}\, e^{+\frac{2\pi i}{h}(p\cdot r)} \tag{8}$$

to be compared with the classical expression (4). One sees easily that if in (8) one could write for:

$$e^{-H/kT} = e^{-V/kT} \cdot e^{+(h^2/8\pi^2 m kT)\Delta} \tag{9}$$

one would get exactly (4). But (9) is *wrong*, because V and Δ are *not commutable*. One comes so to the question to find a general development for the difference $e^{p+q} - e^p \cdot e^q$ if p and q are not commutable. A quite explicit series for this difference is I think not in the mathematical literature.[6] One can of course work it out step by step, and in this way one can derive the final result.

Another way is the following: Call:

$$F = e^{-\beta H} \cdot e^{2\pi i (p\cdot r)/h} \tag{10}$$

[5] E. Wigner, Phys. Rev. **40**, 749, 1932; G. E. Uhlenbeck and L. Gropper, Phys. Rev. **41**, 79, 1932. Lately J. G. Kirkwood (Phys. Rev. **44**, 31, 1933) has derived the same results in a more elegant way. I have not been able to find in the mathematical literature a treatment of series like (2) in connection with the eigenvalue problem. One cannot reduce (2) to a series of the iterated kernels of the to (1) corresponding integral equation, because one cannot develop $e^{-\beta E_n}$ into a series of negative powers of E_n!

[6] Comp. J. E. Campbell, Theory of continuous groups, pp. 54–57 for analogous questions.

where $\beta = 1/kT$. Then F will fulfill the equation:

$$\frac{\partial F}{\partial \beta} = -HF = \frac{h^2}{8\pi^2 m}\Delta F - VF \tag{11}$$

which one might call the to (1) *associated parabolic partial differential equation*, or as I am used to calling it: the *Bloch equation*.[7] To find $F(xyz\, p_x\, p_y\, p_z)$ one has to solve (11) with the *initial condition* $F = e^{2\pi i(p\cdot r)/h}$ for $\beta = 0$ and with the same *boundary conditions* as are imposed on ψ in (1). This can be done by appropriate successive approximation methods. The nicest way is I think as Kirkwood does. He puts:

$$F = we^{2\pi i(p\cdot r)/h + \beta H(p,x)}$$

in which w is a new independent variable (for $\beta = 0, w = 1$) and $H(p, x)$ is the Hamilton function which enters in (3). Developing w in a series of powers of h:

$$w = \sum_\tau h^\tau w_\tau$$

substituting and equating equal powers of h, one can find successively w_τ and so one obtains:

$$S_{qu} = S_{class.}\left[1 - \frac{\pi^2 h^2}{3mk^2T^2}\left\{\Delta V - \frac{1}{2kT}(grad\ V)^2\right\} + \cdots\right] \tag{12}$$

This has just the form as predicted at the end of §2.[8] The higher terms will contain higher powers of h, higher order differential quotients and higher powers and products of the lower ones. These one can calculate successively. Observe that this method only works if the $V(xyz)$ is analytic (or better can be differentiated any number of times). We cannot therefore apply the result to the elastic sphere case; there one has to go back to the exact series of the form (7).

Finally I should like to point out a general connection between the series (2) and eq. (11). Eq. (11), being a generalized conduction of heat equation, one might try to solve with the help of the *Green function* $G(x_i, \xi_i, \beta)$, which is roughly that solution which *a.* for $\beta \to 0$ behaves for x_i near ξ_i as the point source in ξ_i, and *b.* fulfills the boundary con-

[7] F. Bloch, ZS.f. Phys. **74**, 295, 1932.

[8] To see this one has only to put $V/kT = v$, $x/\lambda = \xi$ etc., where λ is again $h/(mkT)^{\frac{1}{2}}$.

ditions. One proves then easily that the series (2) is just $G(x_i, x_i, \beta)$! One has to observe only that according to the definition of G:

$$F = \int_{-\infty}^{+\infty}\int\int\int d\xi_1\, d\xi_2\, d\xi_3\, e^{2\pi i(p-\xi)/h} G(x_i, \xi_i, \beta)$$

Substituting in (6) gives the result according to Fourier's integral theorem.

RANDOM FUNCTIONS

By N. Wiener

By a random function we understand a function which, in a certain way to be specified in what follows, depends upon a formally expressed variable and a usually suppressed parameter of distribution. Such functions occur throughout statistical mechanics. In statistical mechanics a certain possible state of the universe is expressed by a function of one or more variables which give the geometrical coordinates and the time, together with a parameter which singles out the universe considered as one among all possible universes and determines its probability. As a more specific example let us take the path of a particle subject to a Brownian motion and let us consider one coordinate of the particle (let us say the X-coordinate) as a function of the time t. Then for any particular Brownian motion or motion of a particle impelled by the collision of neighboring molecules subject to a thermal agitation, x will be a well-defined function of t. If however, instead of considering the actual traces by a specified particle, we consider all possible courses traced by all possible particles, in addition to the variable t, x will have an argument which singles out the specific Brownian motions in question from all possible Brownian motions. This variable is introduced for purposes of integration; that is, a certain range of this variable measures by its length the probability of the set of Brownian motion it represents, and an integration with respect to this variable yields a probability average of the quantity integrated. Naturally this probability theory is not at all simple and needs to be specified in detail.

It will be seen that the theory of random functions is in essentials a theory of integration in function space. As such it may be subsumed under the general theories of integration of Radon, Daniell, and Wiener. A previous attempt at a theory of integration in function space was due to Gâteaux, but this earlier theory is not a special case of the Daniell-Radon integral. The earlier theory attempted to treat every value $F(t_0)$ of a function $F(t)$ as an independent variable. In such a scheme we find that a succession of regions of positive measure may include one another successively in such a way that there is no element common to all of them. This violates one of the most essential canons

of the Daniell theory. In order to eliminate it, it is necessary that the class of functions over which we are integrating should be in some sense or other compact; that is, that any sequence of functions should contain a subsequence with a limit function. If this is not strictly true with the original class with which we are dealing, it must at any rate be possible to make it true by removing or adjoining a set of functions of arbitrarily small measure. Compactness in the ordinary sense is equivalent to uniform boundedness and equi-continuity. Now, it is a priori obvious that Gâteaux's type of integration must yield us functions which almost never are continuous, and completely breaks down on this point.

The Brownian motion points us a way out of the difficulty. In this motion it is not the position of a particle which is independent of the position at another time, but the travel of a particle between two stated times which is independent of the position at the first time. In crude language, the differentials of the function $x(t)$ correspond to the independent coordinates of a point in space of a finite number of dimensions, and not the values of the function. Physically it is at least reasonable that the Brownian motion of a particle is continuous, and we shall in fact show that the Einstein theory of the Brownian motion permits us to say that in fact it is almost always continuous. We shall show even more, that it is subject to a condition of equi-continuity except for a set of cases whose measure, or what is equivalent, whose probability, we can reduce to as small a value as we want.

The theory of random functions always makes the impression of a much greater degree of artificiality than corresponds to the facts. The reason is that in the theory of Daniell or Lebesgue integration of the sum of a denumerable set of null sets is itself null, while the sum of a set of null sets of the power of the continuum need not be null. This fact makes it extremely desirable to characterize a random function by a denumerable set of conditions, while the characterization of such a function by its value for all its arguments yields a number of characterizations of the power of continuum. If then we attempt to characterize a random function by a set of its values, it is almost necessary for the purposes of our technical development that the original set of values which we take should be denumerable. For example, we may define such a function by its values at the rational or the binary points of the line. On the other hand, we may give up entirely the characterization of such a function by its values and define it by its Fourier coefficients or its coefficients in some other scheme of development. In

any case it will turn out that the function which we have so defined will, except in a set of cases of probability zero (or measure zero, which is the same thing), either be continuous or in some readily specified sense be equivalent to a continuous function. Once this is done, we may define the value of this continuous function uniquely over the whole of its interval of definition. We thus have generalized our initial function defined by a denumerable set of parameters to a function apparently defined by a larger number of parameters. In doing so we have made use of a method which masks the true invariance of the result we have obtained. It is, however, a perfectly legitimate mathematical method to obtain our result in a non-invariantive way and to establish its proper invariance later by specific theorems. This mode of procedure unluckily has the disadvantage of being most non-heuristic and of demanding from the reader the patience to take on faith the necessity of a large amount of material whose justification is only given after the completion of the argument. It will ultimately turn out that the random functions we deal with and the measure that properly belongs to them determine a function-space distinct from that of Hilbert but covariant with it under all unitary transformations which leave Hilbert space invariant. While such a theory may be built up for real Hilbert space, and while this is the case in my previous writings, the present theory of differential space applies to one covariant with complex Hilbert space.

It will be seen that although this note is apparently devoted to integration in a continuous infinity of dimensions, it really forms a chapter in what E. Borel has designated the theory of "denumerable probabilities."[1] The first close approach to the specific problems which concern us here was made by Steinhaus.[2] He considers the series

$$\sum_{1}^{\infty} \pm \frac{1}{n}$$

where the signs $\pm$ represent independent choices, and shows that this series converges in almost every case. The methods of Steinhaus were made into a powerful analytic tool by Paley and Zygmund,[3] more espe-

[1] E. Borel, Les probabilités dénombrables et leurs applications arithmetiques, Rendiconti del Circolo Matematico di Palermo, vol. 27 (1909), pp. 247–271.

[2] Cf. Sur la probabilité de la convergence de series, Studia Mathematica, vol. 2 (1930), pp. 21–39, and earlier papers there referred to.

[3] On some series of functions, Proceedings of the Cambridge Philosophical Society, vol. 26 (1930), pp. 337–357, 458–474; vol. 28, pp. 190–205. Cf. also Notes

cially in the formation of Gegenbeispiele, and the final theory has been employed with great effect by Bohr and Jessen,[4] in the study of the Riemann zeta function and of almost periodic function in the complex domain.

Wiener[5] has developed a theory of random functions in many respects parallel to those discussed by the authors already mentioned, but not identical with them. Since the Wiener theory of random functions is derived from considerations in the theory of the Brownian motion and of statistical mechanics, in which Gaussian distributions play an important rôle, they also play an important rôle in his theory. In this respect, Wiener's theory resembles the Einstein theory of the Brownian motion, to which we shall later show it to be equivalent.[6] Indeed a sharp distinction may be made between Wiener's theory and all others so far mentioned in that, while in both cases a denumerable set of terms are assigned coefficients with a certain distribution, in all other theories these coefficients either have at random the value ± 1, or are distributed at random around the unit circle in the complex plane. In addition to its applicability to statistical mechanics, Wiener's theory possesses a larger degree of symmetry then the alternative theories, or what is the same thing in other words, it has a more extensive group of transformations under which it is invariant. The reason for this is that if a number of terms have independent Gaussian distributions, any linear combination of these terms has itself a Gaussian distribution. On the other hand, if we start with a distribution of the original terms over the values ± 1, or around the unit circle, the distribution of a linear combination of the original terms becomes extremely complex and unmanageable. This property of Gaussian distribution is thus intimately allied with the covariance which the theory of random functions developed here shows with Hilbert space.

By a change of scale, a real Gaussian distribution may be reduced to the distribution of a real parameter uniformly over the interval (0, 1). In the same way, a complex Gaussian distribution may be reduced to

---

on random functions, Paley, Wiener, and Zygmund, Mathematische Zeitschrift, vol. 37 (1933), pp. 647–688, on which the present chapter is largely based.

[4] H. Bohr and B. Jessen, Über die Werteverteilung der Riemannschen Zetafunktion, Acta Mathematica, vol. 54 (1930), pp. 1–35; vol. 58 (1932), pp. 1–55.

[5] Cf. N. Wiener, Generalized harmonic analysis, Acta Mathematica, vol. 55, pp. 214 ff., and the references there given.

[6] A. Einstein, Annalen der Physik, vol. 17 (1905), p. 549 ff.; vol. 19 (1906), p. 371 ff.

the simultaneous and independent uniform distribution of two parameters over an interval of that sort. By this artifice our integration in function space may be carried back to the integration of a function of a denumerable infinity of variables over a cube in its proper space. Such an integration has been discussed by Daniell, Jessen and others.[7] Since however it is probably not familiar to the reader of this note, we shall derive it from the ordinary integration of a function of one variable over an interval of a line by process of mapping. This process of mapping is merely an elaboration of the process by which a square may be mapped on a line in such a way that planar measures are mapped into equal linear measure.

The points of Wiener's space correspond to the formal Fourier series

$$\sum_{-\infty}^{\infty} a_n e^{inx} \tag{1}$$

where each a_n has real and imaginary components all independent and all subject to a Gaussian distribution with the same modulus. We may determine the distribution of all the coefficients a_n by a single parameter, distributed uniformly. Such a distribution may be defined by a Daniell measure. The series (1) almost never converges in the mean, but the series arising from them by formal integration: namely

$$\psi(x, \alpha) = a_0 x + \sum_{1}^{\infty} \left(\frac{a_n e^{inx}}{in} + \frac{a_{-n} e^{-inx}}{-in} \right)$$

almost always converges in the mean. It determines a function which has been shown to be almost always continuous, and indeed almost always to satisfy uniform Lipschitz conditions of every order below $\frac{1}{2}$. The probability that there exists any point for which it satisfies a Lipschitz condition of order exceeding $\frac{1}{2}$ is zero. Accordingly, almost all such functions share the properties of the Weierstrass non-differentiable functions.

If $a < b < c < d$, the correlation between the excursus

$$\psi(a, \alpha) - \psi(b, \alpha)$$

and that $\psi(c, \alpha) - \psi(d, \alpha)$ is zero. Indeed, these quantities are not merely linearly, but completely independent in their distributions. If

[7] P. J. Daniell, Integrals in an infinite number of dimensions, Annals of Mathematics (2), vol. 20 (1919), pp. 281-288. B. Jessen, Bidrag til Integratteorien for Funktioner af uendelig mange Variable, Copenhagen, 1930.

t be taken as the time, $R\psi(t, \alpha)$ be plotted along the x-axis, and $I\psi(t, \alpha)$ along the Y-axis, the function $\psi(t, \alpha)$ will represent a typical Brownian motion of a particle in a fluid, impelled by the molecular agitation of the fluid.

The formal series (1) is transformed by the unitary transformation T into the formal series

$$\sum_{-\infty}^{\infty} a_n T e^{inx},$$

where Te^{inx} is any closed set of normal and orthogonal functions. This may be rearranged into a series

$$\sum_{-\infty}^{\infty} b_n e^{inx}$$

by expanding the functions Te^{inx} in Fourier series. It may be shown that the new coefficients b_n all have completely independent Gaussian distribution for their real and imaginary parts with the same moduli as those of the a_n's.

Let T be any unitary transformation, and let T^n be its n^{th} iterate. If for every functions $f(x)$ and $g(x)$ of L_2,

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} (T^n f(x))\, \bar{g}(x)\, dx = 0$$

then for almost all α,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{0}^{N-1} \varphi\left(\int_{-\pi}^{\pi} T^n f(x)\, d\psi(x, \alpha)\right) = \int_0^1 \varphi\left(\int_{-\pi}^{\pi} f_n(x)\, d\psi(x, \alpha)\right) d\alpha .$$

Here we use Birkhoff's ergodic theorem.

We may easily define the random function $\psi(x, \alpha)$ over $(-\infty, \infty)$ by a change in variable. We may construct the analytic function

$$\int_{-\infty}^{\infty} f(z + x)\, d\psi(x, \alpha)$$

dependent on the parameter z. The problem of the distribution of the zeros of this function is an analytic problem quite analogous to Dr. Jessen's problem of the zeros of an almost periodic function in the complex domain. We show under certain very general restrictions that if $N_A(y_0, y_1)$ is the number of zeros in the rectangle

$$|Rz| < A, \qquad y_0 < Iz < y_1,$$

and

$$\varphi(u) = \text{l.i.m.}_{A\to\infty} \int_{-A}^{A} f(x+iy)\, e^{-iu(x+iy)}\, dx,$$

then

$$N_A(y_0, y_1) \sim \frac{A}{2\pi}\left\{\frac{\int_{-\infty}^{\infty} 2u\,|\varphi(u)|^2 e^{-2uy_0}\,du}{\int_{-\infty}^{\infty} |\varphi(u)|^2 e^{-2uy_0}\,du} - \frac{\int_{-\infty}^{\infty} 2u\,|\varphi(u)|^2 e^{-2uy_1}\,du}{\int_{-\infty}^{\infty} |\varphi(u)|^2 e^{-2uy_1}\,du}\right\}.$$

I wish to emphasize the essential similarity of the problems we discuss and those treated by Bohr and Jessen. The coefficients of Bohr and Jessen are distributed around the unit circle, whereas mine have complex Gaussian distributions. So long as we adhere to one special system of orthogonal functions, one distribution is as natural as the other. However, the distribution which we give is, as we have seen, not tied to any particular set of normal and orthogonal functions.[8]

[8] Cf. Paley and Wiener, Fourier Transforms in the Complex Domain, New York, 1934, Chapters IX and X.

SOME ANALYTICAL PROBLEMS RELATING TO PROBABILITY

By B. Jessen

The problems I am going to talk about are of a purely analytical character but have a certain probability aspect. All they have to do with the calculus of probabilities is that I use the words probability and expectation; but I might just as well omit them, and use the words measure and integral.

From the mathematical point of view we are concerned with applications of the theory of measure and integration in a certain space of an enumerable number of dimensions. This means that from the point of view of probabilities we are concerned with a certain case of enumerable probabilities. Consider the space of all sequences

$$\epsilon_1, \epsilon_2, \epsilon_3, \cdots$$

of complex signs. By a complex sign ϵ I mean a complex number for which $|\epsilon| = 1$. I shall call this space Q_ω. Each of the coordinates describes the unit circle in the complex plane and each one is independent of all others. We may therefore call Q_ω an infinite-dimensional torus space. In this space we have a theory of measure and integration in the Lebesgue sense. The measure of the space itself, that is the total probability, is 1.

The theory of measure and integration in this space has been used by Daniell[1] and Wiener[2] as an example for more general theories. In more recent years it has been studied by Steinhaus,[3] Paley and Zygmund[4] and myself. I have given a detailed exposition of it in my dissertation,[5] a more complete exposition will soon appear in Acta mathematica. The

[1] P. J. Daniell, Integrals in an infinite number of dimensions, Ann. of Math. (2) **20** (1919), 281–288.

[2] N. Wiener, The mean of a functional of arbitrary elements, Ann. of Math. (2) **22** (1921), 66–72.

[3] H. Steinhaus, Sur la probabilité de la convergence de séries I, Studia math. **2** (1930), 21–39. Über die Wahrscheinlichkeit dafür, dass der Konvergenzkreis einer Potenzreihe ihre natürliche Grenze ist, Math. Zeitschr. **31** (1930), 408–416.

[4] R. E. A. C. Paley and A. Zygmund, On some series of functions (1), (2), (3), Proc. Cambridge Philos. Soc. **26** (1930), 337–357; 458–474; **28** (1932), 190–205.

[5] B. Jessen, Bidrag til Integralteorien for Funktioner af uendelig mange Variable, Copenhagen 1930, 1–66.

theory is here developed in analogy to the theory of the n-dimensional torus space. In the recent book of Kolmogoroff on the foundations of the calculus of probabilities[6] we find a very similar general exposition of the theory of probabilities involving an infinite number of choices. He does not consider precisely our case but the case of real coordinates each of which has a certain distribution; this case was also treated by Daniell.

If we do not consider sequences of complex signs but sequences of real signs ± 1 we get a classical case which has been much studied, mostly by means of dyadic numbers. The theory of Paley and Zygmund has been developed for both cases. It is important to emphasize that my own applications lead to the case of complex signs. I shall now show by a number of examples how the theory applies to analytical problems. I first give two applications of Steinhaus.

(i) Consider a sequence $r_1, r_2, r_3, \cdots$ of non-negative numbers and form all series $\sum_1^\infty \epsilon_n r_n$ where the absolute values of the terms are the given numbers and the signs are determined by chance. What is the probability that it shall converge? Or in other words: What is the measure of the set of points in Q_ω for which the series converges? The answer is that the probability is 1 if $\sum_1^\infty r_n^2 < \infty$ and 0 if $\sum_1^\infty r_n^2 = \infty$. In the case of real signs ± 1 the result is due to Rademacher and Kolmogoroff; the theorem is a special case of a much more general theorem due to Kolmogoroff.

(ii) Consider again a sequence $r_1, r_2, r_3, \cdots$ and suppose that $\overline{\lim}\ \sqrt[n]{r_n} = 1$. Consider all power series $f(z) = \sum_1^\infty \epsilon_n r_n z^n$. They all have the radius of convergence 1. What is the probability for such a series to be continuable? Or in other words: What is the measure of the set of points in Q_ω for which $f(z)$ is continuable? The answer is that the probability is always 0. The proof is easy. The case of real signs ± 1 is more difficult; it has been treated by Paley and Zygmund.

(iii) These are both examples of the nought-and-one law in the calculus of probabilities. A number of theorems of this kind have been given by Paley and Zygmund. A large part of their theorems cluster

[6] A. Kolmogoroff, Grundbegriffe der Wahrscheinlichkeitsrechnung. Ergebnisse der Mathematik 2, Heft 3, Berlin 1933, 1–62.

about (ii); they have given results concerning order and type of almost all power series; they have also considered Fourier series and Dirichlet series.

I now come to some of my own applications.

(iv) Consider a sequence $r_1, r_2, r_3, \cdots$ of non-negative numbers and a sequence of real numbers $\lambda_1, \lambda_2, \lambda_3, \cdots$ which are all different, and consider formally the quite general Dirichlet series $f(s) = \sum_1^\infty \epsilon_n r_n e^{\lambda_n s}$. I can prove: Let $\alpha < \sigma < \beta$ be the largest interval in which the series $\sum_1^\infty r_n^2 e^{2\lambda_n \sigma}$ converges; then our series are almost all uniformly convergent in any bounded domain inside the strip $\alpha < \sigma < \beta$ and the functions $f(s)$ are almost all not continuable across the borderlines of the strip. This result is also true for real signs ± 1; it generalises a theorem of Paley and Zygmund.

I have also proved that the functions $f(s)$ are almost all almost periodic in $[\alpha, \beta]$ in the generalised sense of Besicovitch. The proof of this result requires the use of the ergodic theorem of Birkhoff in the space Q_ω. The reason why this theorem can be used is that if we replace s by $s + i\tau$ for a real τ the class of functions $f(s)$ is transformed into itself. Let T_τ denote the transformation of Q_ω into itself corresponding to the translation $s + i\tau$; then the transformations T_τ are simply translations in Q_ω in the "direction" $(\lambda_1, \lambda_2, \lambda_3, \cdots)$ and they satisfy the conditions for the application of the ergodic theorem.[7] This argument does not hold for real signs ± 1. The theorem contains a recent result of Carlson.[8]

(v) Consider the Zeta-function $\zeta(s) = \prod_1^\infty (1 - p_n^{-s})^{-1}$ and form all functions $f(s) = \prod_1^\infty (1 - \epsilon_n p_n^{-s})^{-1}$. These functions are all regular and $\neq 0$ for $\sigma > 1$. It is easy to prove that the functions are almost all regular and $\neq 0$ in the larger halfplane $\sigma > \frac{1}{2}$ and also that they are almost all not continuable across the line $\sigma = \frac{1}{2}$. This is true also for real signs ± 1.

[7] I apply the theorem in the form in which it was proved by Khintchine. A. Khintchine, Zu Birkhoffs Lösung des Ergodenproblems, Math. Ann. **107** (1933), 485–488.

[8] F. Carlson, Contributions à la théorie des séries de Dirichlet. Note III, Arkiv för Mat. **23A** N:o 19 (1933), 1–8.

Now Bohr and I have proved[9] that if $N_a(T)$ denotes the number of a-points of $\zeta(s)$ in the rectangle $(\frac{1}{2} <) \alpha < \sigma < \beta$, $|t| < T$ the limit $H_a = \lim N_a(T)/2T$ will exist. I can show that if we denote by n_a the number of a-points for $f(s)$ in $\alpha < \sigma < \beta$, $|t| < \frac{1}{2}$ the limit H_a is nothing but the expectation for n_a. The proof is obtained by a passage to the limit from the corresponding theorem for the functions $\zeta_n(s) = \prod_1^n (1 - p_\nu^{-s})^{-1}$ and $f_n(s) = \prod_1^n (1 - \epsilon_\nu p_\nu^{-s})^{-1}$; instead of the ergodic theorem of Birkhoff, we here apply the theorem of Kronecker-Weyl. Again this argument does not generalise to the case of real signs ± 1.

An observation which probably everybody has made who has worked in this field is: If we want to show the existence of a function with a pathological property it is often the easiest to show that in an appropriate sense almost all functions have this property.

[9] H. Bohr und B. Jessen, Über die Werteverteilung der Riemannschen Zetafunktion I, II, Acta math. **54** (1930), 1–35; **58** (1932), 1–55.

PRINCIPLES OF PROBABILITY IN NATURAL SCIENCE

By Felix Bernstein

The following four typical cases may be considered:

1) Probability of Mendelian heredity

2) Probability of composition of a Mendelian population

3) Probability of death as a function of age

4) Probability of crossing-over and geometry of the chromosome.

(1) If a male which inherited A from one parent, a from the other parent i.e. a male Aa is crossed with a female AA, sperms A and a in equal numbers are produced and meet eggs of type A under equal chances, the condition of arbitrary choice being fulfilled in fertilization, the sperms moving in rather ergodic ways.

In most cases the sperms A and a do not differ in size, shape or quality, the genes they carry being in "silent" stage.

The condition of the Laplace urn is fulfilled in an ideal way.

The cases of equal probability are furnished by Nature herself, which produces enormous masses of cells not distinct in appearance but in hidden qualities not influencing the drawing from the urn.

This production of enormous masses of biological entities of the same or nearly the same quality is analogous to the mass-production of electrons, atoms, etc. and is one of the fundamentals of the successful approach of the natural phenomena by the calculus of probability. As long as we do not see through the way of formation of those uniform masses, we cannot hope to understand the principles of probability in Nature completely.

What we can understand is, this situation being granted, how it works out with the aid of ergodic mechanisms to produce the results we observe.

(2) In populations starting with an initial distribution of certain hereditary types, we can follow the mixture of inbreeding without selection and with selection.

J. B. S. Haldane and R. A. Fisher have shown how the ingenious ideas of Charles Darwin can be developed and justified by exact calculations of great mathematical interest. Non-linear difference equations play an important rôle.

In the theory of human blood-groups a point of attack has been found

for the problem of human races by means of calculus of probability. It has been found that we have to use the surface of Steiner and its representation on the plan by the Weierstrassparameter as the mathematical tool. (Ref. Table of formula 1.)

The observations of the blood-groups in a special population define a point of the Steiner surface.

The Parameter of Weierstrass corresponding to this point gives the proportions of the pure Mendelian races composing the population. As the Steiner surface is of genus zero they can be determined rationally.

The mixture of two different populations represented by two points of the surface with different weights will be effected in determining the centre of gravity of the two corresponding points in the plan of the parameters and then going back to the point of the surface corresponding to the centre of gravity in the parameter plan.

A question of principle arises if observations do not give a point lying exactly on the surface. As the number of observations is finite, this has to be expected always. What is the best determination of a point on the surface? To answer this we have to assume that the plan of the parameters is a plan of equal likelihood, that the rule of Bayes gives at least a best choice of the point on the Steiner surface determined by two other surfaces of 4th degree having one point in common with the surface. (Table 4.)

But the assumption is only acceptable if we have no reason to suppose a trend in the racial composition of the population. Such a trend is effected by geographical variation, because the *BB* race is spreading from an innerasiatic centre, so that *BB* is increased.

Neglecting arbitrarily what we know we get the condition of equal likelihood. This case illustrates clearly that the scheme of cases of equal likelihood of Laplace is very often nothing else but a deliberate unwillingness to go into further detail in order to get a first approximation to the facts. This is a very common situation in biological statistics. In criticising Heisenberg's principle of indeterminedness Ludwik Silberstein holds the same point of view in regard to the principles of the quantum theory.

(3) In vital statistics the same situation arises even in a much more expressed sense. Practically we know in a human population the utter inhomogeneity of the cases and knowing the names of the human beings involved we could have any desired information. But we neglect everything except age and sex in making, for instance, a table of normal mortality.

Making a table of impaired lives we change arbitrarily again the groups of equal probability in taking into account classifications of impairment by diseases. Laplace's viewpoint is definitely changed in considering cases as of equal probability, not if we do not know about their difference, but if we do not like to know about their well-known difference.

Gompertz has derived a mathematical law of mortality as a function of age:

$$\mu_x' = a\mu_x$$

μ_x designating the instantaneous mortality.

To make clearer what is involved I make the following experiment: A certain number of Italian spaghetti were cooked and stretched and pasted in parallel order on a board, the ends attached at the side so that a certain tension was reached. If the spaghetti dried they broke. The relative number depends on the relation of elasticity to tension. This breaking is analogous to the occurrence of death, elasticity corresponding to the state of aging, tension to the stress of environment, mortality of the spaghetti is proportional to decrease in elasticity with the loss of water. The ergodic scheme will give a perfect explanation of this connection.

T. H. Morgan showed that the hereditary qualities of a plant or an animal are due to small invisible entities called chromosomes. They are certain strings of stainable material forming kernels of the cell. Each species has a definite number of different chromosomes — $c_1 \cdots c_n$ occurring in a single set in the germ cells and in a double set $c_1 \cdots c_n$; $\bar{c}_1 \cdots \bar{c}_n$ in the body cells. The body cells forming the germ cells show juxtaposition of paired chromosomes $c, c_1', \cdots c_n', c_n'$ and in each pair c_1, c_1' there are random breaks of the chromosome pairs. By exchange of the corresponding parts, new pairs $d_1, d_2, \cdots d_n, d_n'$ are formed (crossing over). Then separation at random in two germ cells follows, so that we have two sets $e_1 \cdots e_n$; $e_1' \cdots e_n'$ each belonging to a different germ cell and selected at random, so that $e_1 = d_1$ or d_1' with chance $\frac{1}{2}$. The interesting part is the break of the paired chromosomes at random, because they give the means to define and distinguish two genes A B so far as a break and exchange occurs between A and B. The distance AB can be determined too by the probability of such a break. As AB/C cannot be distinguished from $A.C.B$ for determination of order are needed two basic points AB. Then the order CD becomes different from DC in all experiments separating A

from B and C from D: (i.e. $A/BC/D$ and $A/BD/C$). In consequence order type of the geometry can be secured by mere qualitative experiment. If there is nearly the same survival in the different classes, the probability CAB of separation of AB gives a measure of distance AB by assuming independent probability for separation of AB and BC. We have the following formulae:

$$C_{AC} = C_{AB}(1 - C_{AB}) + C_{BC}(1 - C_{AB})$$

$$1 - 2C_{AC} = (1 - 2C_{AB})(1 - 2C_{BC})$$

It follows that const. log $(C1 - 2C_{AB})$ is a measure of eculidian distance in a linear geometry. Generalization of such a geometry of breaks to higher dimensions is easy. I show a box filled with clay which by drying up will have breaks separating two arbitrary spots A and B in this space.

I discuss now the principles of such a geometry. In our geometry we have not a system of coördinates in which every coördinate has a determined rational value but exp distance* is a frequency determined with distribution, connection depending on the number N of observations. Such a coördinate (to have a name) we may call a Pascal coördinate, shorter: a *pasc*, instead of a Descartes coördinate or a *desc*. The point in a pasc-geometry is not an invisible, but his determination or better definition depends on the number N of observations of breaks and he may be dissolved in more points by further observations. This is really the situation in the case of a hereditary gene. Pasc-geometry tends to a sort of desc-geometry if N becomes infinite. But in the limit only *almost all* points of the supposed desc-geometry can be interpreted. Indeed, given a series of approximating pasc-geometries to a desc-geometry D, the relationship of the m becomes the simplest, if we alter a little the definition of distance, defining distance AB as a mean value of the separations observed between A and B. This mean value will not be affected if we suppress a set of intervals in each pasc-geometry with the sum tending to zero if N becomes infinite. Substituting for a desc-geometry such a limiting pasc-geometry, such a theorem as Birkhoff's ergodic theorem would guarantee an absolute affirmation of objective truth of ergodic character of the ways of a particle.

I used such a pasc coördinate in a paper on celestial mechanics in 1911, proving that there is no mean motion of earth and Venus in the sense of Lagrange. In genetics we cannot go to the limit of N because

if we continue an experiment too long undiscovered mutations of genes in other genes would come in and change our fundamental probabilities of crossing over. We have a fundamental *incoherence* of our determinations with time. A similar situation occurs in the experiment of interference of polarized rays of light in optics.

That the incoherence even of the most accurate periodic processes makes it impossible to identify points and distances completely after a certain lapse of time may create a general situation to make impossible to use real limit geometry for more than a first approximation of the observations, so that we have to rely on pasc N-geometry with some great N in the mechanics of Nature. Pasc or desc measurement in space and time are based on identification of wave lengths of light.

Appendix

Let an arbitrary point in the space of three dimensions be designated by the homogenous coördinates

$$x_1, x_2, x_3, x_4$$

Then the equation of the Steiner surface, as given by Weierstrass in the parametrical form, may be written

$$\begin{aligned} \rho x_1 &= (\xi_1 - \xi_2 - \xi_3)^2 \\ \rho x_2 &= (\xi_2 - \xi_3 - \xi_1)^2 \\ \rho x_3 &= (\xi_3 - \xi_1 - \xi_2)^2 \\ \rho x_4 &= (\xi_1 + \xi_2 + \xi_3) \end{aligned} \tag{1}$$

The equation of the surface in the irrational form is immediately derived

$$\sqrt{x_1} + \sqrt{x_2} + \sqrt{x_3} + \sqrt{x_4} = 0 \quad \text{(Cayley)} \tag{2}$$

The surface being of genus zero, it is possible to express the parameters $\xi_1\xi_2\xi_3$ as polynomials in x_1, x_2, x_3, x_4. In order to get those expressions the equations (1) may be transformed so that in the three first ones the douple p products appear. To this effect we introduce the new variables

$$\begin{aligned} 4X_1 &= -x_1 - x_2 + x_3 + x_4 \\ 4X_2 &= -x_2 - x_3 + x_1 + x_4 \\ 4X_3 &= -x_3 - x_1 + x_2 + x_4 \\ 4X_4 &= +x_1 + x_2 + x_3 + x_4 \end{aligned} \tag{3}$$

Evidently we have

$$\begin{aligned} \rho X_1 &= 2\xi_1\xi_2 \\ \rho X_2 &= 2\xi_2\xi_3 \\ \rho X_3 &= 2\xi_3\xi_1 \\ \rho X_4 &= \xi_1^2 + \xi_2^2 + \xi_3^2 \end{aligned} \tag{4}$$

The identity

$$\xi_1 : \xi_2 : \xi_3 = \xi_1\xi_2 \cdot \xi_1\xi_3 : \xi_2\xi_3 \cdot \xi_2\xi_1 : \xi_3\xi_1 \cdot \xi_3\xi_2 \tag{5}$$

leads to the proportion

$$\xi_1 : \xi_2 : \xi_3 = X_1X_3 : X_2 \cdot X_1 : X_3X_2 \tag{6}$$

Designating with X the sum

$$X_1X_3 + X_2X_1 + X_3X_2$$

we get

$$\begin{aligned} \xi_1/\xi_1 + \xi_2 + \xi_3 = X_1X_3/X\,, \qquad \xi_2/\xi_1 + \xi_2 + \xi_3 = X_2X_1/X\,, \\ \xi_3/\xi_1 + \xi_2 + \xi_3 = X_3X_2/X \end{aligned} \tag{7}$$

In similar way we get the proportion

$$\xi_1 : \xi_2 : \xi_3 = X_2^{-1} : X_3^{-1} : X_1^{-1} \tag{8}$$

By using the identity

$$\rho^2 X = 4\xi_1\xi_2\xi_3\,(\xi_1 + \xi_2 + \xi_3) \tag{9}$$

we get in combination with the last line of (1)

$$\frac{\xi_1}{\xi_1 + \xi_2 + \xi_3} = \frac{X}{2X_2\,(X_1 + X_2 + X_3 + X_4)}, \cdots \tag{10}$$

and by adding the three expressions the Steiner surface in rational form:

$$1 = X^2/2X_1X_2X_3\,(X_1 + X_2 + X_3 + X_4)$$

In the same way the Cayley equation can be considered as the sum of the three determinations of ξ_1, ξ_2, ξ_3 in irrational form. As Landsteiner discovered, each population offers the observation of not more than four different groups, the blood of each group allowing transfusion of blood without reaction inside of each group. Heredity of the blood-groups was studied by Ottenberg v. Dungern and Hirschfeld. As I

found, the groups can be described in terms of the hereditary composition of each individual by two of a set of three genes

$$A^*, \quad B^*, \quad R^*$$

one gene inherited from the father, the other from the mother. The formula for each individual is therefore a combination of the six pairs of genes

$$A^*A^*, \quad A^*R^*, \quad B^*B^*, \quad B^*R^*, \quad R^*R^*, \quad A^*B^*.$$

But AA and A^R fuse to A and BB and BR fuse to B. The group O has the composition RR and the group AB has the composition AB.

Designating with p, q, r the frequencies of the genes and the frequencies of the groups with the letters of the groups we have the problem, to determine the p, q, r if the O, A, B, AB are given by observation. The latter are expressed by the former by the expressions

$$(11) \quad A = p^2 + 2pr, \quad B = q^2 + 2qr, \quad O = r^2 \quad AB = 2pq.$$

It is obvious that the problem is overdetermined.

The observed bloodgroups have to fulfill a relation which is a consequence of the hypothesis of the hereditary scheme, which is tested by the expected relation.

This relation is the equation

$$(12) \qquad \sqrt{A+O} + \sqrt{B+O} - \sqrt{O} - 1 = D = O$$

which is the equation of the surface of Steiner. To get the form (1), (2), (10) we put

$$(13) \qquad O + B{:}O + A{:}O{:}O + A + B + AB = x_1{:}x_2{:}x_3{:}x_4$$

and on the other side we put

$$(14) \qquad \xi_1 = \frac{p}{2}, \quad \xi_2 = \frac{q}{2}, \quad \xi_3 = \frac{p+q}{2} + r.$$

The parameter p, q, r can be considered as the proportions of three pure Mendelian races which generate by homogenous mixture the population observed. The decomposition in pure races corresponds, therefore, to the determination of the parameter of Weierstrass of the surface.

A more general problem arises if we take into consideration the fact that a finite number of observations cannot determine exactly a point

of the surface. We have to expect a chance deviation and we have to determine the most probable value of p, q, r.

The distribution being given by

$$(15)\quad W = \frac{(O+A+B+AB)}{O!\,A!\,B!\,AB!}\,(r^2)^O\,(p^2+2pr)^A\,(q^2)^B\,(2pq)^{AB}$$

we have to resolve the equations

$$(16)\qquad \frac{\partial W}{\partial p} = 0 \qquad \frac{\partial W}{\partial q} = 0 \qquad (r = 1-p-q)$$

or

$$(17)\qquad \frac{O}{r^2} = \frac{A}{(p+r)^2 - r^2} - \left(\frac{B}{(q+r)^2 - r^2} - \frac{AB}{2pq}\right)\frac{q}{r}$$

$$= \frac{B}{(q+r)^2 - r^2} - \left(\frac{A}{(q+r)^2 - r^2} - \frac{AB}{2pq}\right)\frac{p}{r}$$

Approximate solutions are given by the expressions

$$(18)\quad p = (1-\sqrt{B+O})\left(1+\frac{D}{2}\right), \quad q = (1-\sqrt{A+O})\left(1+\frac{D}{2}\right)$$

SUR LES MOYENNES ARITHMÉTIQUES DES SÉRIES-NOYAUX DES DÉVELOPPEMENTS EN SÉRIES D'HERMITE ET DE LAGUERRE ET SUR CELLES DE CES SÉRIES-NOYAUX DERIVÉES TERME A TERME

PAR ERVAND KOGBETLIANTZ*

INTRODUCTION

On connaît bien que les propriétés de la série-noyau

$$(1) \qquad \varphi_0(x)\varphi_0(u) + \varphi_1(x)\varphi_1(u) + \cdots + \varphi_n(x)\varphi_n(u) + \cdots$$

du développement suivant les fonctions d'un système orthogonal normé d'une fonction $f(x)$

$$f(x) \sim \sum_{n=0}^{\infty} \varphi_n(x) \int_a^b f(u)\varphi_n(u)\, du$$

jouent un rôle prépondérant dans l'étude des conditions suffisantes de convergence ou de sommabilité de ce développement. Ceci explique bien la nécessité de l'étude des moyennes arithmétiques des deux séries-noyaux

$$(H) \qquad 0 \sim \sum_0^{\infty} \frac{H_n(x)H_n(u)}{2^n n!\sqrt{\pi}} \qquad (u \neq x)$$

et

$$(L) \qquad 0 \sim \sum_0^{\infty} \frac{\Gamma(n+1)}{\Gamma(n+\alpha+1)} L_n^{(\alpha)}(x) L_n^{(\alpha)}(u) \qquad \begin{pmatrix} \alpha > -1 \\ x \neq u \end{pmatrix}$$

relatives aux systèmes de polynômes d'Hermite $H_n(x)$ et de ceux de Laguerre $L_n^{(\alpha)}(x)$, dont les premiers sont orthogonaux avec le poids e^{-x^2} dans l'intervalle $(-\infty, +\infty)$ et les deuxièmes—avec le poids $x^\alpha e^{-x}$—dans l'intervalle (O, ∞), en supposant $\alpha > -1$.

Ces séries (H) et (L) représentent zéro partout pourvu que l'on ait $x \neq u$, ce résultat étant démontré dans ce qui suit. Cela veut dire

* Docteur ès Sciences (Paris), Professeur d'Analyse à Téhéran (Perse).

PHYSICS LIBRARY

qu'étant divergentes elles sont néanmoins sommables par la méthode des moyennes arithmétiques et sommables avec une somme nulle. Observons en passant que l'étude des séries-noyaux des différents systèmes orthogonaux a abouti jusqu'à maintenant au même résultat qui paraît être général: une série-noyau du type (1)—nous parlons des séries-noyaux étudiées—représente pour $u \neq x$ zéro, même si elle diverge, car elle est sommable par tel ou autre procédé de sommation régulier avec une somme nulle. Il serait extrêmement important de former un système de fonctions orthogonales tel que la série-noyau correspondante à ce système soit sommable pour $u \neq x$, mais avec une somme non-nulle partout, car un tel système—s'il peut exister—présenterait sûrement des propriétés bien curieuses en ce qui concerne les développements suivant les fonctions de ce système.

Dans ce mémoire nous établissons pour les moyennes arithmétiques $H_n^{(\delta)}(x, u)$ et $L_n^{(\delta)}(x, u)$ des séries (H) et (L) les inégalités fondamentales suivantes:

$$(AH) \qquad H_n^{(\delta)}(x, u) = O\left(\frac{e^{\frac{x^2+u^2}{2}}}{|u - x|^{\delta+1}\sqrt{n^\delta}}\right)$$

$$(AL) \qquad L_n^{(\delta)}(x, u) = O\left(\frac{e^{\frac{x+u}{2}}(ux)^{-\alpha/2-\frac{1}{4}}}{|\sqrt{u} - \sqrt{x}|^{\delta+1}\sqrt{n^\delta}}\right)$$

où $\delta \geqq 0, \alpha > -1$.

Non seulement les séries (H) et (L), mais aussi les séries qu'on obtient en les dérivant terme à terme par rapport à x, sont très importantes. En particulier, leurs moyennes arithmétiques $\frac{\partial}{\partial x} H_n^{(\delta)}(x, u)$ et $\frac{\partial}{\partial x} L_n^{(\alpha)}(x, u)$ jouent un rôle important dahs l'étude des conditions suffisantes imposées à l'allure de $f(x)$ et qui permettent de tirer du developpement en série la valeur du saut de $f(x)$ en un point $x = x_0$, où les deux valeurs $f(x_0 \pm 0)$ existent mais sont inégales. C'est pourquoi nous étudions dans ce mémoire les expressions $\frac{\partial}{\partial x} H_n^{(\delta)}(x, u)$ et $\frac{\partial}{\partial x} L_n^{(\delta)}(x, u)$ également et démontrons entre autres les inégalités suivantes:

$$(BH) \qquad \frac{\partial H_n^{(\delta)}(x, u)}{\partial x} = O\left(\frac{e^{\frac{x^2+u^2}{2}}\sqrt{n}}{|x - u|^{\delta+1}\sqrt{n^\delta}}\right)$$

$$(BL) \qquad \frac{\partial L_n^{(\delta)}(x,u)}{\partial x} = O\left(\frac{e^{\frac{x+u}{2}}(ux)^{-\alpha/2-\frac{1}{4}}\sqrt{n}}{\sqrt{x}\,|\sqrt{u}-\sqrt{x}|^{\delta+1}\sqrt{n^\delta}}\right).$$

Toutes ces inégalités (A) et (B) exigent l'emploi de la méthode dite "du col" pour pouvoir les démontrer dans l'intervalle de variation de u: $u_0 < |u| < \eta\sqrt{n}$ pour $H_n^{(\delta)}$ et $\frac{\partial}{\partial x} H_n^{(\delta)}$ et dans celui $u_0 < u < \eta n$ pour $L_n^{(\delta)}$ et $\frac{\partial}{\partial x} L_n^{(\delta)}$, η et u_0 étant deux quantités positives fixes choisies d'avance dont η est aussi petite qu'on veut et u_0 aussi grande qu'on veut. Pour les valeurs finies de u on peut se servir des formules approchées formées d'après le théorème général de M. Maitland Wright.

Pour $\eta\sqrt{n} \leqq |u| \leqq e\sqrt{n}$ (polynômes d'Hermite) et $\eta n \leqq u \leqq e^2 n$ (polynômes de Laguerre) on ramène les moyennes d'ordres non-entiers à celles d'ordres entiers, ces dernières pouvant être étudiées plus facilement. Enfin, pour $|u| > e\sqrt{n}$ et $u > e^2 n$ respectivement les inégalités en question se laissent démontrer à partir des expressions explicites exprimant $H_n^{(\delta)}$, $L_n^{(\delta)}$ et leurs dérivées par rapport à x à l'aide de polynômes de Laguerre. C'est pourquoi nous commençons par donner une nouvelle démonstration de l'inégalité fondamentale relative au polynôme de Laguerre

$$(2) \qquad L_n^{(\alpha)}(x) = O\left[e^{\frac{x}{2}} n^{\alpha/2-\frac{1}{4}} x^{-\alpha/2-\frac{1}{4}}\right],$$

que nous avons démontrée déjà dans un mémoire relatif aux séries d'Hermite.[1] Cette nouvelle démonstration est liée à la méthode du "col" et avant de l'exposer nous donnons dans le §1 un aperçu sur cette méthode si importante.

§1. MÉTHODE DU COL

Cette méthode sert à l'évaluation approchée des intégrales définies du type

$$J = \frac{1}{2\pi i}\int_{(C)} F(z)\,dz$$

étendue à une courbe fermée (C) entourant l'origine ou un autre point fixé d'avance dans le plan de la variable complexe $z = x + iy$. On

[1] E. KOGBETLIANTZ. *"Recherches sur la sommabilité des séries d'Hermite."* Annales de l'Ecole Normale Supér. (3), XLIX, 1932, Paris, pp. 137–221.

connaît bien comment on l'applique pour obtenir les formules approchées (voir p. ex. le beau mémoire de MM. M. Plancherel et W. Rotach dans les Commentarii Mathem. Helvetici, t. I, pp. 227–254).

Nous allons exposer dans ce qui suit la méthode générale à suivre pour obtenir la meilleure *borne supérieure* pour $|J|$.

Posons à cet effet

$$F(z) = G(x, y)e^{iH(x,y)}$$

$\log G(x, y)$ et $H(x, y)$ étant deux fonctions harmoniques. Les dérivées partielles de G et H sont liées par les relations classiques

$$G'_y + GH'_x = 0 \qquad G'_x - GH'_y = 0$$

$$G^{-1}G''_{x^2} = H'^2_y + H''_{xy} \qquad G^{-1}G''_{y^2} = H'^2_x - H''_{xy}$$

$$G^{-1}G''_{xy} = H''_{y^2} - H'_xH'_y .$$

On constate que les extrema des deux fonctions $G(x, y)$ et $H(x, y)$ coïncident, car $G'_x = G'_y = 0$ entraine $H'_x = H'_y = 0$ et vice-versa. Les valeurs des dérivées secondes de $G(x, y)$ en un point (x_0, y_0) où s'annulent à la fois G'_x et G'_y, sont données par les relations suivantes:

$$G''_{x_0^2} = G_0H''_{x_0y_0}, \qquad G''_{x_0y_0} = G_0H''_{y_0^2}, \qquad G''_{y_0^2} = -GH''_{x_0y_0}$$

d'où l'on tire

$$G''_{x_0y_0} - G''_{x_0^2}G''_{y_0^2} = G_0^2(H''^2_{y_0^2} + H''^2_{x_0y_0}) > O .$$

Cela prouve que les extrema de $G(x, y)$ sont des cols, c'est-à-dire des points hyperboliques de la surface $\zeta = G(x, y)$, d'où le nom de la méthode. On en profite en faisant passer le chemin d'intégration (C) par ces cols et par les vallées de la surface $\zeta = G(x, y)$. Dans ces conditions les maxima de $G(x, y)$ sur (C) sont atteints précisément aux cols. Soit z_0 un col pour $\zeta = G(x, y)$. Ce point z_0 est un col aussi pour la surface $\zeta = H(x, y)$, donc les affixes des cols sont racines de l'équation

$$F'(z) = 0 .$$

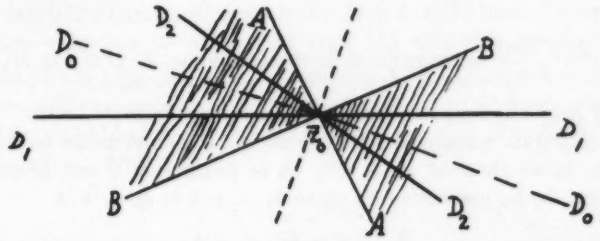
En ce point z_0 les directions asymptotiques sur la surface $\zeta = G(x, y)$ sont données par l'équation

$$G''_{x^2}\cos^2\varphi + 2G''_{xy}\sin\varphi\cos\varphi + G''_{y^2}\sin^2\varphi = 0$$

tandis que celles sur la surface $\zeta = H(x, y)$ sont déterminées par l'équation

$$H''_{x^2}\cos^2\varphi + 2H''_{xy}\sin\varphi\cos\varphi + H''_{y^2}\sin^2\varphi = 0 .$$

Tenant compte du fait qu'au col z_0 on a $G'_x = G'_y = H'_x = H'_y = 0$ et par conséquent $G''_{x^2} = GH''_{xy}$, $G''_{xy} = GH''_{y^2}$, $G''_{y^2} = -GH''_{xy}$ ainsi que $H''_{x^2} + H''_{y^2} = O$, on constate que les directions asymptotiques sur la surface $\zeta = H(x, y)$ sont bissectrices des angles droits formés par les directions asymptotiques sur la surface $\zeta = G(x, y)$, ce qui veut dire qu'au voisinage



immédiat du col z_0 les axes des vallées pour $\zeta = G(x, y)$ coïncident avec les directions asymptotiques sur $\zeta = H(x, y)$. Considérons une droite passant par le point z_0 (figurée en projection sur le plan xOy) et comprise dans la région hachurée sur la figure qui représente les vallées de la surface $\zeta = G(x, y)$.

Quelle que soit sa position dans l'angle droit $Az_0\,B$ la valeur de la fonction G sur cette droite passe par un maximum au point z_0, tandis que celle de la fonction H y passe soit par un minimum, soit par un maximum suivant que la droite en question est de part ou d'autre de la direction asymptotique (axe de la vallée) figurée par la ligne $\overline{D_0D_0}$. Sur cette ligne $\overline{D_0D_0}$ elle même la dérivée seconde de $H(x, y)$ s'annule au point z_0. Cette droite $\overline{D_0D_0}$, qui est la ligne de la plus grande pente pour la surface $\zeta = G(x, y)$ au voisinage du point z_0, est déterminée par l'angle ψ_0 qu'elle fait avec Ox et dont la tangente est donnée par l'une des deux expressions suivantes

$$H''_{x^2} \operatorname{tang} \psi_0 = H''_{xy} \pm \sqrt{H''^2_{xy} + H''^2_{x^2}},$$

les valeurs des dérivées secondes de H étant celles relatives au point z_0. En effet cette tangente doit vérifier l'équation classique

$$G''_{xy}\xi^2 + (G''_{x^2} - G''_{y^2})\xi - G''_{xy} = 0$$

et l'on choisit le signe devant le radical, en posant $z = z_0 + te^{i\psi_0}$ et en exigeant que l'on ait au point z_0 donc pour $t = 0$

$$G''_{t^2} < 0\,.$$

On trouve ainsi l'expression définitive suivante pour $\tan \psi_0$

$$H''_{x^2} \tan \psi_0 = H''_{xy} + \sqrt{H''^2_{xy} + H''^2_{x^2}}$$

ainsi que

$$H''_{x^2} + H''_{xy} \tan 2\psi_0 = 0$$

donc

$$e^{2i\psi_0} \frac{F''(z_0)}{F(z_0)} = \text{quantité réelle et négative} = -\sqrt{H''^2_{x^2} + H''^2_{xy}}$$

ce qui peut aussi servir pour déterminer ψ_0.

Considérons maintenant l'intégrale de $F(z)dz$ étendue au segment $\overline{PQ}$ de la droite $z = z_0 + te^{i\psi}$, où le paramètre ψ est constant et différent de ψ_0, mais compris entre $\psi_0 - \pi/4$ et $\psi_0 + \pi/4$

$$0 < |\psi - \psi_0| < \pi/4 .$$

Le col correspond à la valeur nulle du paramètre variable t et soient $-l$ et $+m$ les valeurs de t relatives aux extrémités P et Q du segment. On a

$$\frac{1}{2\pi i} \int_{\overline{PQ}} F(z)dz = \frac{e^{i(H_0+\psi)}}{2\pi i} \int_{-l}^{+m} Ge^{i(H-H_0)}\, dt$$

où $H_0 = H(0)$ est la valeur de $H(t)$ au col. L'argument $H - H_0$ des fonctions cos et sin s'annule au col et y possède un extremum.

La fonction positive $G(t)$ atteint au col, pour $t = 0$, son maximum, donc pourvu que $G'(t)$ et $H'(t)$ ne s'annulent qu'une seule fois sur le segment $\overline{PQ}$, on peut écrire

$$\left| \int_{-l}^{+m} Ge^{i(H-H_0)}\, dt \right| < 2G_0(\lambda + \mu) ,$$

$-\lambda$ et μ étant les deux racines de l'équation

$$|H(t) - H(0)| = \pi$$

appartenant à l'intervalle $(-l, m)$. Or, vue que $\psi \neq \psi_0$, $H'(0) = 0$ tandis que $H''(0) \neq 0$, on a:

$$|H(t) - H(0)| = \frac{t^2}{2} |H''(0)| \cdot \left| 1 + \frac{tH'''(\vartheta t)}{3H''(0)} \right| = \pi .$$

En choisissant l et m suffisamment petits pour avoir $H''(t) \neq 0$ tant que $-l \leqq t \leqq m$, on constate que les valeurs approchées de λ et de μ

sont égales à $\sqrt{2\pi}\,|\,H''(0)\,|^{-\frac{1}{2}}$ et plus précisément λ et μ sont inférieures à $\sqrt{2\pi}N^{-\frac{1}{2}}$, où $N = \min |\,H''(t)\,|$, t variant de $-\lambda$ à μ. Ainsi, on a le résultat

$$J = \frac{1}{2\pi i}\int_{\overline{PQ}} F(z)\,dz = O(G_0 N^{-\frac{1}{2}}).$$

Quant aux contributions provenant des arcs de (C) joignant entre eux différents segments de droite du type $\overline{PQ}$ elles sont généralement négligeables vis à vis des termes du type $G_0 N^{-\frac{1}{2}}$ à condition de choisir judicieusement ces arcs. On parvient ainsi, s'il y a plusieurs cols et si l'on désigne par G_k et N_k les valeurs que prennent $G(x, y)$ et N au *k-ième* col, au résultat suivant:

$$J = O\left(\sum_k G_k N_k^{-\frac{1}{2}}\right).$$

Il est évident que le choix judicieux des directions ψ_k à donner aux segments $\overline{P_k Q_k}$ joue un grand rôle car il permet de réduire considérablement les calculs.

Les applications de cette méthode conduisent souvent aux intégrales définies réelles du type

$$j = \int_a^b \vartheta(x)\cos[f(x)]dx$$

où $\vartheta(x)$ et $f(x)$ sont monôtones dans (a, b) et $f''(x) \neq 0$ dans l'intervalle ouvert (a, b). Le second théorème de moyenne permet de démontrer facilement le lemme suivant qui nous sera utile dans la suite.

LEMME: *Soit* $M = \max \vartheta(x) = \vartheta(a)$ *ou* $= \vartheta(b)$ *suivant que* $\vartheta'(x)$ *est négative ou positive dans* (a, b). *Soit aussi* $c = a$ *ou* $c = b$ *suivant que l'on a* $f''(x) > 0$ *ou* $f''(x) < 0$ *dans* (a, b). *Dans l'hypothèse* $f'(x) > 0$ *dans l'intervalle ouvert* (a, b) *on a*

$$|\,j\,| < M\left\{\left[\left(\frac{f'(c)}{f''(c)}\right)^2 + \frac{2\pi}{|\,f''(c)\,|}\right]^{\frac{1}{2}} - \left|\frac{f'(c)}{f''(c)}\right|\right\} = Mh$$

si pour $f''(x) > 0$ *est remplie la condition* $f'''(x) < 0$ *dans l'intervalle* $(a, a + h)$ *ou bien celle* $f'''(x) > 0$ *dans l'intervalle* $(b - h, b)$ *dans le cas* $f''(x) < 0$. *Enfin, si l'on a* $f'(c) = 0$ *la condition du signe imposée à* $f'''(x)$ *n'étant pas remplie, on a tout de même*

$$|\,j\,| < M \cdot \sqrt{\frac{6\pi}{|\,f''(c)\,|}} = Mh^*$$

pourvu que l'on ait dans $(a, a + h^*)$ [*ou* $(b - h^*, b)$]

$$| f'''(x) |^2 < \frac{2}{3\pi} \cdot | f''(c) |^3 .$$

Nous n'insistons pas sur la preuve de ce lemme, sa démonstration étant facile à reconstituer.

§2. INÉGALITÉ POUR LE POLYNÔME DE LAGUERRE

Dans ce paragraphe nous appliquons la méthode du col aux polynômes de Laguerre $L_n^{(\alpha)}(x)$ définis par

$$(1 - z)^{-(\alpha+1)} e^{-\frac{xz}{1-z}} = \sum_0^\infty z^n L_n^{(\alpha)}(x) .$$

L'expression

$$n!\, x^\alpha e^{-x} L_n^{(\alpha)}(x) = \frac{d^n}{dx^n} (x^{n+\alpha} e^{-x})$$

permet d'écrire

$$2\pi i x^\alpha e^{-x} L_n^{(\alpha)}(x) = \int_{(C)} z^{n+\alpha} e^{-z} (z - x)^{-(n+1)}\, dz = \int_{(C)} F(z)\, dz,$$

le chemin d'intégration (C) entourant le point $z = x$.

L'équation $F'(z) = 0$, où

$$F'(z) = F(z) \left[\frac{n + \alpha}{z} - 1 - \frac{n + 1}{z - x} \right],$$

donne les deux cols

$$z = \frac{x + \alpha - 1}{2} \pm \frac{i}{2} \sqrt{4(n + \alpha)x - (x + \alpha - 1)^2}$$

qui sont complexes et conjugués tant que

$$(\sqrt{n + \alpha} - \sqrt{n + 1})^2 < x < (\sqrt{n + \alpha} + \sqrt{n + 1})^2 .$$

Nous supposons d'abord $x \leqq n(1 - \epsilon)$, $\epsilon > 0$, et soient

$$z = \xi + i\eta = re^{i\theta}, \qquad z - x = ue^{i(\pi - \varphi)}$$

ainsi que

$$F(z) = G(\xi, \eta) e^{uH(\xi, \eta)}$$

d'où

$$G \equiv e^{-\xi} r^{n+\alpha} u^{-(n+1)},$$

$$H \equiv -\eta + (n+\alpha)\theta + (n+1)\varphi - (n+1)\pi.$$

On trouve facilement

$$H''_{\xi^2} = 2\eta[r^{-4}(n+\alpha)\xi + u^{-4}(n+1)(x-\xi)],$$

$$H''_{\xi\eta} = r^{-4}(n+\alpha)(\eta^2 - \xi^2) + u^{-4}(n+1)[(x-\xi)^2 - \eta^2].$$

Au col $z_0 = \xi_0 + i\eta_0 = \frac{1}{2}\{x + \alpha - 1 + i\sqrt{4(n+\alpha)x - (x+\alpha-1)^2}\}$ on a

$$|z_0| = r_0 = \sqrt{(n+\alpha)x}, \qquad u_0 = |z_0 - x| = \sqrt{(n+1)x}$$

donc au col:

$$x^2(n+\alpha)(n+1)H''_{\xi^2} = \eta_0[(2n+\alpha+1)x - (\alpha-1)^2],$$

$$x^2(n+\alpha)(n+1)H''_{\xi\eta} = 2(1-\alpha)\eta_0^2$$

et par conséquent

$$\tan\psi_0 = \frac{H''_{\xi\eta} + \sqrt{H''^2_{\xi^2} + H''^2_{\xi\eta}}}{H''_{\xi\eta}}$$

$$= 1 + \frac{1-\alpha}{\sqrt{nx}}\sqrt{1 - \frac{x}{2(2n+\alpha+1)}}\left\{1 + O\left(\frac{1}{\sqrt{nx}}\right)\right\}.$$

Pour $\alpha = 1$ on a exactement $\psi_0 = \frac{\pi}{4}$ tandis que pour $\alpha \neq 1$ on a

$$\psi_0 = \frac{\pi}{4} + \frac{1-\alpha}{2\sqrt{nx}}\sqrt{1 - \frac{x}{2(2n+\alpha+1)}}\left\{1 + O\left(\frac{1}{\sqrt{nx}}\right)\right\}.$$

En désignant la valeur de l'angle polaire θ au col par θ_0, on a aussi pour $x < 2(2n+\alpha+1)$

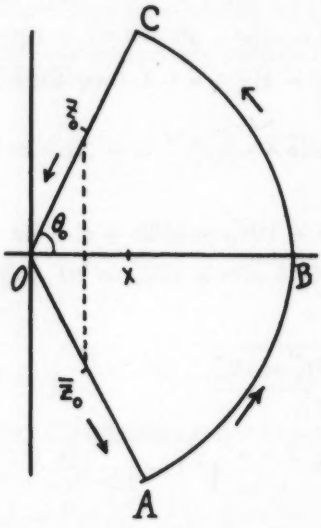
$$|\tan(\theta_0 - \psi_0)| = \left|1 - \sqrt{\frac{x}{n}} + O(n^{-1})\right| < 1$$

donc $|\theta_0 - \psi_0| < \pi/4$ ce qui prouve que la direction $\theta = \theta_0$ passe dans la vallée du col. En choisissant comme chemin d'intégration (C) les deux demi-droites issues de l'origine et passant par les cols z_0 et $\bar{z}_0$

complètées par un arc de cercle $\overset{\frown}{ABC}$ de centre 0 et de rayon $R > x$ on peut écrire:

$$2\pi i x^{\alpha} e^{-x} L_n^{(\alpha)}(x) = \int_{\overline{OA}} + \int_{\overset{\frown}{ABC}} + \int_{\overline{CO}} F(z)\, dz = J_1 + J_2 + J_3 .$$

Or,



$$|J_2| \leqq 2R^{\alpha} e^{-R\cos\theta_0} \int_0^{\theta_0} \frac{d\theta}{\left(1 - \frac{x}{R}\right)^{n+1}} .$$

Etant donné que $\cos\theta_0 = \dfrac{x + \alpha - 1}{2\sqrt{(n+\alpha)x}} > 0$ (c'est évident pour $\alpha \geqq 1$, tandis que pour $|\alpha| < 1$ nous supposons $x > 2(1 - \alpha)$), on constate que J_2 s'évanouit pour $R \to \infty$

$$\lim_{R=\infty} J_2 = 0$$

et l'on trouve définitivement, vu que $z = te^{i\theta_0}$ sur $\overline{OC}$

$$\pi x^{\alpha} e^{-x} L_n^{(\alpha)}(x) = R\left\{ e^{i(\pi/2+\theta_0)} \int_0^{\infty} F(te^{i\theta_0}) dt \right\}$$

les cols correspondant à $t = t_0 = \sqrt{(n+\alpha)x}$. Posons en outre

$$z = te^{i\theta_0} = x + \rho e^{i\omega}$$

d'où

$$\omega = \arctan\left(\frac{t\sin\theta_0}{t\cos\theta_0 - x}\right), \rho = \sqrt{x^2 - 2xt\cos\theta_0 + t^2}$$

ainsi que

$$e^{i(\pi/2+\theta_0)}F(te^{i\theta_0}) = G(t)e^{iH(t)}$$

d'où

$$G(t) = e^{-t\cos\theta_0}\rho^{-(n+1)}t^{n+\alpha},$$

$$H(t) = (n+\alpha+1)\theta_0 + \pi/2 - t\sin\theta_0 - (n+1)\omega.$$

On trouve facilement la valeur de la dérivée $H'(t)$

$$\rho^2 H'(t) = -\sin\theta_0(t - t_0)\left(t + \frac{n+1-x}{\sqrt{n+\alpha}}\sqrt{x}\right)$$

donc, vu que par hypothèse $x \leqq n(1-\epsilon)$

$$H'(t) \gtrless 0 \quad \text{suivant que} \quad t \gtrless t_0$$

et cela quelque soit t dans l'intervalle $(0, \infty)$.

De même, en formant la dérivée $G'(t)$, on trouve:

$$\rho^2 t\frac{G'(t)}{G(t)} = -(t - t_0)[At^2 + Bt + C]$$

avec $A = \cos\theta_0$, $C = xt_0$ et

$$B = \frac{x - \alpha + 1}{2} - 2x\cos^2\theta_0.$$

Le discriminant du trinôme $At^2 + Bt + C$ est négatif pourvu que l'on ait

$$2(1-\alpha) < x < (2\sqrt{2}+1)(n+\alpha) + \frac{(\sqrt{2}+1)(1-\alpha)}{2\sqrt{2}+1} + O(n^{-1})$$

ce qui prouve que dans l'hypothèse $2(1-\alpha) \leqq x \leqq n(1-\epsilon)$.

$G'(t)$ ne s'annule dans $(0, \infty)$ qu'une seule fois pour $t = t_0$ et $G'(t) \gtrless 0$ suivant que $t \gtrless t_0$. Ainsi, les deux fonctions $G(t)$ et $H(t)$ atteignent leurs maxima pour $t = t_0$ et décroissent ensuite constamment.

En calculant $H''(t)$ et $H'''(t)$, on trouve

$$H''(t) = -2(n+1)\, x \sin\theta_0\, \rho^{-4}\, (t - x\cos\theta_0)$$

et

$$H'''(t) = 2x(n+1)\, \sin\theta_0\, \rho^{-6}\, (3\rho^2 - 4x^2 \sin^2\theta_0)\,.$$

Par conséquent, $H''(t)$ ne s'annule qu'une seule fois dans $(0, \infty)$ à savoir pour $t = x\cos\theta_0$ et, vu que $x \leqq n(1-\epsilon)$, on constate que $x\cos\theta_0$ est inférieur à $\frac{1}{2}\, t_0$:

$$x\cos\theta_0 = x\,\frac{x+\alpha-1}{2\sqrt{(n+\alpha)\,x}} < \frac{x(n+\alpha)}{2\sqrt{(n+\alpha)\,x}} = \tfrac{1}{2}\, t_0.$$

L'hypothèse $x \leqq n(1-\epsilon)$ entraine également $\theta_0 > \pi/3$. En effet

$$(x+\alpha-1)^2[\tan^2\pi/3 - \tan^2\theta_0] = 4\{x^2 - (n-\alpha+1)x + (\alpha-1)^2\} < 0\,.$$

Or, les deux racines du trinôme sont

$$x_1 = \frac{(\alpha-1)^2}{n}\,[1 + O(1/n)] \quad \text{et} \quad x_2 = n - \alpha + 1 + O(1/n)$$

donc en vertu de l'hypothèse $2(1-\alpha) < x \leqq n(1-\epsilon)$ on a constamment $x_1 < x < x_2$. Ceci entraine le fait que $H'''(t)$ est positive dans l'intervalle $[t_0 - O(\sqrt[4]{nx}),\ t_0]$. En effet, on a $H'''(t) < 0$ tant que $3\rho^2 < 4x^2 \sin^2\theta_0$, c'est à dire tant que

$$-\frac{2x}{\sqrt{3}} \sin\left(\theta_0 - \frac{\pi}{3}\right) < t < \frac{2x}{\sqrt{3}} \sin\left(\theta_0 + \frac{\pi}{3}\right)$$

où

$$\frac{2'}{\sqrt{3}} \sin\left(\theta_0 + \frac{\pi}{3}\right) < \frac{2}{\sqrt{3}} \sin\frac{2\pi}{3} = 1\,,$$

et $H'''(t)$ est positive pour $t \geqq x$.

Or,

$$t_0 - O(\sqrt[4]{nx}) = \sqrt{nx} \cdot \left[1 + O\left(\frac{1}{\sqrt[4]{nx}}\right)\right]$$

donc

$$\frac{x}{t_0 - O(\sqrt[4]{nx})} = \sqrt{\frac{x}{n}}\left\{1 + O\left(\frac{1}{\sqrt[4]{nx}}\right)\right\} < 1$$

ce qui prouve que $H'''(t) > 0$ pour $t \geqq t_0 - O(\sqrt[4]{nx})$.

Par contre $H'''(t)$ est négative au voisinage du zéro et plus précisément $H'''(t) < 0$ pour

$$0 \leqq t \leqq \frac{2x}{\sqrt{3}} \sin\left(\frac{\pi}{2} + \frac{\pi}{3}\right) = \frac{x}{\sqrt{3}}.$$

Ceci posé, étudions l'intégrale définie

$$\pi x^{\alpha} e^{-x} L_n^{(\alpha)}(x) = \int_0^{\infty} G(t) \cos\left[H(t)\right] dt$$

en la décomposant en trois:

$$\int_0^{\infty} = \int_0^{x \cos \theta_0} + \int_{x \cos \theta_0}^{t_0} + \int_{t_0}^{\infty} = j_1 + j_2 + j_3.$$

Dans l'intervalle $(0, x \cos \theta_0)$ nous avons $H' > 0$, $G' > 0$ et $H'' > 0$. D'après le lemme on a le droit de conclure:

$$| j_1 | < G(x \cos \theta_0) \left\{ \left[\left(\frac{H_0'}{H_0''} \right)^2 + \frac{2\pi}{H_0''} \right]^{\frac{1}{2}} - \frac{H_0'}{H_0''} \right\} = h_1 \, G(x \cos \theta_0)$$

car $H'''(t) < 0$ au voisinage de l'origine $t = 0$. En effet, on a

$$H_0' = H'(0) = \frac{n + 1 - x}{x} \sin \theta_0$$

et

$$H_0'' = H''(0) = \frac{(x + \alpha - 1)(n + 1) \sin \theta_0}{x^2 \sqrt{(n + \alpha)\, x}}$$

ce qui prouve que le terme $H_0'^{3} \, H_0''^{-2}$ est de l'ordre de nx tandis que celui $2\pi H_0''^{-1}$ est de l'ordre de $x^{\frac{1}{2}} \, n^{-\frac{1}{2}}$. Par conséquent:

$$h_1 = \frac{H_0'}{H_0''} \left\{ \sqrt{1 + 2\pi H_0'' \, H_0'^{-3}} - 1 \right\}$$

$$= \frac{\pi}{H_0'} \left[1 + O\left(\frac{1}{n} \sqrt{\frac{x}{n}} \right) \right] = \frac{\pi x}{(n + 1 - x) \sin \theta_0} \left[1 + O\left(\frac{1}{n} \right) \right].$$

Vu que $h_1 = O\left(\frac{x}{n}\right) < \frac{x}{\sqrt{3}}$ on constate que $H'''(t) < 0$ dans l'intervalle $(0, h_1)$ ce qui assure la validité du lemme dans ce cas:

$$| j_1 | < \frac{\pi x \, G(x \cos \theta_0)}{(n + 1 - x) \sin \theta_0} \left[1 + O(n^{-1}) \right] = O\left[G\left(\frac{t_0}{2} \right) \right]$$

car $2x \cos \theta_0 < t_0$. Or,

$$G\left(\frac{t_0}{2}\right) = e^{-\frac{x+\alpha-1}{4}} \left(\frac{t_0}{2}\right)^{n+\alpha} \left[x^2 - xt_0 \cos\theta_0 + \frac{t_0^2}{4}\right]^{-\frac{n+1}{2}}$$

$$= O\left\{e^{-\frac{x}{4}} [x(n+\alpha)]^{\frac{n+\alpha}{2}} [4x^2 - 2x(x+\alpha-1) + x(n+\alpha)]^{-\frac{n+1}{2}}\right\}$$

$$= O\left\{e^{-\frac{5x}{4}} (nx)^{\frac{\alpha-1}{2}}\right\}, \text{ car } \left[1 + \frac{2(x-\alpha+1)}{n+\alpha}\right]^{-\frac{n+1}{2}} = O(e^{-x})$$

donc

$$|j_1| = O\left[e^{-\frac{5x}{4}} (nx)^{\frac{\alpha-1}{2}}\right].$$

Passons à l'intervalle $(x \cos \theta_0, t_0)$. Dans cet intervalle on a $G' > 0$, $H' > 0$, mais $H'' < 0$. Le lemme pour être valable exige que l'on ait $H''' > 0$ dans l'intervalle $(t_0 - h_2, t_0)$, où à cause de $H'(t_0) = 0$ on a

$$h_2 = \sqrt{\frac{2\pi}{|H''(t_0)|}} = \left\{\frac{2\pi(n+\alpha)^2 \sqrt{x(n+\alpha)}}{(n+1)(2n+\alpha+1-x)\sin\theta_0}\right\}^{\frac{1}{4}} = O(\sqrt[4]{nx}).$$

Or, plus haut on a vu déjà que $H'''(t)$ est positive dans l'intervalle $[t_0 - O(\sqrt[4]{nx}), t_0]$ et le lemme nous donne:

$$|j_2| < h_2 G(t_0) < G_0 \sqrt[4]{(n+\alpha)x} \sqrt{\frac{\pi}{\sin\theta_0\left(1 - \frac{x}{2n+\alpha+1}\right)}},$$

où

$$G_0 = G(t_0) = e^{-t_0 \cos\theta_0} \cdot \rho_0^{-(n+1)} t_0^{\lambda+\alpha} = e^{-\frac{x}{2}} (nx)^{\frac{\alpha-1}{2}} [1 + o(1)].$$

Ainsi

$$|j_2| < e^{-\frac{x}{2}} (nx)^{\alpha/2 - \frac{1}{4}} \left\{\frac{\sqrt{\pi}}{\sqrt{\left(1 - \frac{x}{2n+\alpha+1}\right)\sin\theta_0}} + o(1)\right\}.$$

Enfin, pour la troisième intégrale j_3 on a $H' < 0$, $H'' < 0$ et $H''' > 0$. Posons $H_1 \equiv -H$ de sorte que:

$$j_3 = \int_{t_0}^{\infty} G(t) \cos[H_1(t)]\, dt$$

les dérivées H_1', H_1'' étant positives et H_1''' — négative dans l'intervalle (t_0, ∞). Le lemme est applicable et l'on a:

$$|j_3| < G_0 \sqrt{\frac{2\pi}{|H''(t_0)|}} = h_2 G_0 .$$

En additionannt ces résultats, nous obtenons pour $2(1-\alpha) < x < n(1-\epsilon)$ l'inégalité

$$\pi e^{-x} x^{\alpha} |L_n^{(\alpha)}(x)| < \frac{2\sqrt{\pi}\, e^{-\frac{x}{2}} (nx)^{\alpha/2-\frac{1}{4}} [1 + o(1)]}{\sqrt{\sin\theta_0 \left(1 - \frac{x}{2n+\alpha+1}\right)}}$$

ou bien encore vu que $\theta_0 > \frac{\pi}{3}$ et $x < n$:

$$L_n^{(\alpha)}(x) = O\left(e^{\frac{x}{2}} n^{\alpha/2-\frac{1}{4}} x^{-\alpha/2-\frac{1}{4}}\right). \tag{2}$$

La limitation $x > 2(1-\alpha)$ est sans importance car la formule approchée de Fejér pour le polynôme $L_n^{(\alpha)}(x)$ prouve que (2) est valable pour toute valeur finie de x. D'autre part l'inégalité[2]

$$n!\, |L_n^{(\alpha)}(x)| \leqq (2x)^n \qquad (x \geqq n+\alpha) \tag{3}$$

valable pour $x \geqq n + \alpha$ entraîne celle (2) pour $x \geqq 7.4n$. En effet, on a pour $x \geqq n + \alpha$ grâce à (3)

$$e^{-\frac{x}{2}} x^{\alpha/2+\frac{1}{4}} n^{\frac{1}{4}-\alpha/2} |L_n^{(\alpha)}(x)| \leqq \frac{e^{-\frac{x}{2}}}{\sqrt{2\pi}} \left(\frac{2ex}{n}\right)^n \left(\frac{x}{n}\right)^{\alpha/2+\frac{1}{4}} [1 + o(1)]$$

et il est facile de prouver que le second membre reste inférieur à $(7.4)^{\alpha/2+\frac{1}{4}} \cdot (2\pi)^{-\frac{1}{2}}$, si $x \geqq 7.4n$. Par conséquent, il suffit d'établir (2) dans l'intervalle $n(1-\epsilon) \leqq x \leqq 7.4n$ pour achever la preuve de sa validité uniforme dans $(0, \infty)$.

Or, pour x compris dans l'intervalle $(c_1 n, c_2 n)$, où c_1 et c_2 sont deux nombres positifs fixes, l'inégalité (2) se réduit à

$$L_n^{(\alpha)}(x) = O\left(x^{-\frac{1}{2}} e^{\frac{x}{2}}\right) \tag{4}$$

et nous allons déduire (4) de l'inégalité

$$L_n^{(-\frac{1}{2})}(x) = O\left(n^{-\frac{1}{2}} e^{\frac{x}{2}}\right) \tag{5}$$

qui n'est autre chose que l'inégalité bien connue[3]

$$H_{2n}(\sqrt{x}) = (-1)^n 4^n n!\, L_n^{(-\frac{1}{2})}(x) = O\left(e^{\frac{x}{2}} n^{-\frac{1}{4}} \sqrt{2^{2n}(2n)!}\right).$$

[2] E. Kogbetliantz. L.c., p. 158.
[3] Ibidem, p. 153, formule (29).

La relation générale[4]

$$\Gamma(\alpha-\beta)\,\Gamma(n+\beta+1)\,L_n^{(\alpha)}(x) = \Gamma(n+\alpha+1)\,x^{-\alpha}\int_0^x (x-u)^{\alpha-\beta-1}u^{\beta}\,L_n^{(\beta)}(u)\,du \tag{6}$$

valable pour $\alpha > \beta$ donne pour $\alpha > -\frac{1}{2}$, en y posant $\beta = -\frac{1}{2}$:

$$L_n^{(\alpha)}(x) = O\left\{e^{\frac{x}{2}}(n/x)^{\alpha}\int_0^x e^{-\frac{x-u}{2}}(x-u)^{\alpha-\frac{1}{2}}\frac{du}{\sqrt{u}}\right\} = O\left[x^{-\frac{1}{2}}e^{\frac{x}{2}}\right],$$

car $c_1 < x/n < c_2$ par hypothèse. Pour $\alpha \leqq -\frac{1}{2}$ on peut employer la relation (6) avec $\beta = -m$, si l'on a $1 - m > \alpha > -m$ avec $m = [m] \geqq 1$. A cet effet, démontrons d'abord (4) pour $\alpha = -m$. L'inégalité (4) a été déjà démontrée pour $\alpha = \frac{1}{2} - m$[5] et l'on en déduit:

$$\begin{aligned} L_n^{(-m)}(x) &= \sum_{i=0}^{n}\frac{\Gamma(n-i-\frac{1}{2})}{\Gamma(n-i+1)\,\Gamma(-\frac{1}{2})}\,L_i^{(\frac{1}{2}-m)}(x) \\ &= O\left\{\sum_{i=0}^{n}(i+1)^{-\frac{3}{2}}\left|L_i^{(\frac{1}{2}-m)}(x)\right|\right\} \\ &= O\left\{x^{-\frac{1}{2}}e^{\frac{x}{2}}\sum_{1}^{\infty}i^{-\frac{3}{2}}\right\} = O\left[x^{-\frac{1}{2}}e^{\frac{x}{2}}\right]. \end{aligned}$$

En outre, pour $x \to 0$ on a

$$m!\,(-1)^m x^{-m}\,L_n^{(-m)}(x) = 1 + o(1)$$

ce qui prouve que l'on peut écrire pour $\alpha > -m$

$$L_n^{(\alpha)}(x) = \frac{\Gamma(n+\alpha+1)\,x^{-\alpha}}{\Gamma(m+\alpha)\,\Gamma(n-m+1)}\int_0^x (x-u)^{\alpha+m-1}u^{-m}\,L_n^{(-m)}(u)\,du\,.$$

En décomposant l'intervalle de l'intégration $(0, x)$ en trois $(0, \eta)$, $(\eta, x-\epsilon x)$, $(x-\epsilon x, x)$, on trouve pour $c_1 n < x < c_2 n$:

$$\frac{x^{-\alpha}\,\Gamma(n+\alpha+1)}{\Gamma(m+\alpha)\,(n-m)!}\int_0^{\eta} = O(n^{m+\alpha}\,x^{m-1}) = o\left(x^{-\frac{1}{2}}e^{\frac{x}{2}}\right),$$

$$\frac{x^{-\alpha}\,\Gamma(n+\alpha+1)}{\Gamma(m+\alpha)\,(n-m)!}\int_{\eta}^{x(1-\epsilon)} = O\left(x^{-\frac{1}{2}}e^{\frac{x}{2}}\,x^{2m+\alpha+\frac{1}{2}}e^{-\frac{\epsilon x}{2}}\right) = o\left(x^{-\frac{1}{2}}\cdot e^{\frac{x}{2}}\right)$$

tandis que

[4] Ibidem, p. 156, formule (32).
[5] Ibidem, p. 158.

$$\frac{x^{-\alpha}\,\Gamma(n+\alpha+1)}{\Gamma(m+\alpha)\,(n-m)\,!}\int_{x(1-\epsilon)}^{x}$$

$$= O\left\{\frac{n^{m+\alpha}}{\sqrt{x}}\int_{1-\epsilon}^{1} e^{\frac{ux}{2}}\,(1-u)^{m+\alpha-1}\,du\right\} = O\left(\frac{e^{\frac{x}{2}}}{\sqrt{x}}\right)$$

ce qui achève la preuve de l'inégalité (2) valable, quelque soit α, dans l'intervalle $0 < x < \infty$.

§3. INÉGALITÉS FONDAMENTALES POUR LA MOYENNE $H_n^{(\delta)}(x, u)$ DE LA SÉRIE-NOYAU (H) ET POUR SA DÉRIVÉE PARTIELLE PAR RAPPORT À x

Le but de ce paragraphe sont les deux inégalités:

(AH) $$H_n^{(\delta)}(x, u) = O\left\{\frac{e^{\frac{u^2+x^2}{2}}}{|u-x|^{\delta+1}\sqrt{n^\delta}}\right\}$$

(BH) $$\frac{\partial H_n^{(\delta)}(x, u)}{\partial x} = O\left\{\frac{e^{\frac{u^2+x^2}{2}}\sqrt{n}}{|u-x|^{\delta+1}\sqrt{n^\delta}}\right\}$$

la variable x restant dans un intervalle fini, $|x| \leqq a$, et celle u variant de $-\infty$ à $+\infty$.

La relation

$$\sigma_n^{(\delta)} = A_n^{(\delta)}\,H_n^{(\delta)}(x, u) = \frac{1}{\sqrt{\pi}}\sum_{m=0}^{n}(-1)^m\,L_m^{\left(-\frac{1}{2}\right)}(s^2)\,L_{n-m}^{\left(\delta+\frac{1}{2}\right)}(d^2)\,,$$

où l'on a posé $(x-u)^2 = 2d^2$ et $(x+u)^2 = 2s^2$, permet de démontrer l'inégalité (AH) à l'aide de celle (3) §2 pourvu que l'on ait $|u| \geqq e\sqrt{n}$.[6]

Etant donné que

$$\frac{dL_n^{(\alpha)}(x)}{dx} = -L_{n-1}^{(\alpha+1)}(x)$$

on trouve pour $\frac{\partial}{\partial x}H_n^{(\delta)}(x, u)$ la relation analogue

$$\frac{\partial\sigma_n^{(\delta}}{\partial x} = A_n^{(\delta)}\frac{\partial}{\partial x}H_n^{(\delta)} = \frac{1}{\sqrt{\pi}}\sum_{m=0}^{n}(-1)^m\left[(u-x)L_m^{\left(-\frac{1}{2}\right)}(s^2)\,L_{n-m-1}^{\left(\delta+\frac{3}{2}\right)}(d^2)\right.$$

$$\left. - (u+x)L_{m-1}^{\left(\frac{1}{2}\right)}(s^2)\,L_{n-m}^{\left(\delta+\frac{1}{2}\right)}(d^2)\right]$$

[6] L.c., pp. 172–173.

où comme d'ordinaire

$$A_n^{(\delta)} = \frac{\Gamma(n+\delta+1)}{\Gamma(\delta+1)n!} \sim \frac{n^\delta}{\Gamma(\delta+1)} \qquad (n \to \infty).$$

A l'aide de (3) et en appliquant le même raisonnement que l'on vient de citer au sujet de $\sigma_n^{(\delta)}$, on déduit pour $|u| \geqq e\sqrt{n}$ l'inégalité (BH).

Ainsi la preuve est faite pour $|u| \geqq e\sqrt{n}$. Considérons maintenant les deux intervalles $\eta\sqrt{n} \leqq |u| \leqq e\sqrt{n}$, η étant une quantité positive aussi petite qu'on veut mais fixe. Etant donné que x est fini, les inégalités à démontrer se réduisent à celles

$$(A_1) \qquad \sigma_n^{(\delta)} = A_n^{(\delta)} H_n^{(\delta)}(x,u) = O\Big(|u|^{-1} e^{\frac{u^2}{2}}\Big),$$

$$(B_1) \qquad \tau_n^{(\delta)} = \frac{\partial}{\partial x}\sigma_n^{(\delta)} = A_n^{(\delta)} \frac{\partial}{\partial x} H_n^{(\delta)}(x,u) = O\Big(e^{\frac{u^2}{2}}\Big).$$

Supposons d'abord $\delta = [\delta] \geqq 0$. L'inégalité (A_1) a été établie déjà[7] dans l'hypothèse $\delta = [\delta]$. On peut établir à l'aide du même raisonnement celle (B_1). Soit d'abord $\delta = 0$. Alors, un calcul facile donne

$$\tau_n^{(0)} = \frac{\partial}{\partial x} H_n^{(0)} = 2^{-(n+1)} (n!)^{-1} (u-x)^{-2} \{2(u-x)[nH_{n+1}(u) H_{n-1}(x) - (n+1) H_n(x) H_n(u)] + H_{n+1}(x) H_n(u) - H_n(x) H_{n+1}(u)\}.$$

Vu que x est fini, on a

$$H_n(x) = O\Big[n^{-\frac{1}{4}} \sqrt{2^n n!}\Big]$$

tandis que

$$H_n(u) = O\Big[e^{\frac{u^2}{2}} n^{-\frac{1}{4}} \sqrt{2^n n!}\Big]$$

et par conséquent pour $|u-x|\sqrt{n} \geqq 1$:

$$\tau_n^{(0)} = O\Big[e^{\frac{u^2}{2}} d^{-1} \sqrt{n}\Big(1 + \frac{1}{d\sqrt{n}}\Big)\Big] = O\Big(e^{\frac{u^2}{2}}\Big),$$

car $d\sqrt{2n} \geqq 1$ et $\sqrt{n} = O(d)$ de même que $d = O(\sqrt{n})$.

Supposons maintenant que (B_1) a été démontrée pour

$$\delta = 0, 1, 2, \cdots k-1$$

[7] L.c., pp. 172-174.

et vérifions qu'alors elle est valable aussi pour $\delta = k$. On connaît pour $\sigma_n^{(\delta)}$, $\delta = [\delta]$ la relation de recurrence.[8]

$$(7) \quad \sigma_n^{(k)} = [2(u-x)]^{-(k+1)}\, G_n^{(k+1)} + \sum_{m=1}^{r} c_{km}\, (u-x)^{-2m}\, \sigma_n^{(k-m)},$$

où r est le plus grand entier contenu dans $\dfrac{k+1}{2}$ et

$$2^n\, n!\, G_n^{(k+1)}\, \sqrt{\pi} = \sum_{0}^{k+1} (-1)^i \binom{k+1}{i} H_{n+k-i+1}(x)\; H_{n+i}(u).$$

En dérivant (7) par rapport à x on trouve

$$\tau_n^{(k)} = 2(k+1)\; [2(u-x)]^{-(k+2)}\; G_n^{(k+1)} + [2(u-x)]^{-(k+1)}\; \frac{\partial}{\partial x} G_n^{(k+1)}$$

$$+ \sum_{m=1}^{r} c_{km} (u-x)^{-2m-1}\; [(u-x)\; \tau_n^{(k-m)} + 2m\; \sigma_n^{(k-m)}].$$

Vu que par hypothèse

$$\tau_n^{(k-m)} = O\left(e^{\frac{u^2}{2}}\right) \qquad (1 \leqq m \leqq k)$$

ainsi que dans l'intervalle considéré:

$$\sigma_n^{(k-m)} = O\left(|\, u\,|^{-1}\, e^{\frac{u^2}{2}}\right), \qquad G_n^{(k+1)} = O\left(e^{\frac{u^2}{2}}\; n^{\frac{k}{2}}\right)$$

et enfin

$$-\frac{\partial}{\partial x} G_n^{(k+1)}$$

$$= 2^{-n}(n!)^{-1}\, \pi^{-\frac{1}{2}} \sum_{0}^{k+1} (-1)^i\, 2(n+k-i+1)\, H_{n+k-1}(x)\, H_{n+i}(u)$$

$$= O\left\{2^{-n}\,(n!)^{-1}\, \sqrt{n} \sum_{0}^{k+1} e^{\frac{u^2}{2}}\; 2^{n+\frac{k}{2}}\; \sqrt{n+k-i)\,!\,(n+i)\,!}\right\}$$

$$= O\left\{e^{\frac{u^2}{2}}\;\; n^{-n}\;\; \sum_{0}^{k+1} n^{\frac{1}{2}(n+k-i+n+i+1)}\right\} = O\left[e^{\frac{u^2}{2}}\;\; n^{\frac{k+1}{2}}\right],$$

[8] L.c., p. 175.

on en déduit l'inégalité (B_1) pour $\delta = k$. Ainsi, la preuve de (A_1) et de (B_1) pour $\delta = [\delta]$ est faite. Soit maintenant $\delta = k - f$, où $k = [\delta + 1] \geqq 0$ et $0 < f < 1$. On a

$$\sigma_n^{(\delta)} = \sum_0^n A_m^{-(f+1)}\, \sigma_{n-m}^{(k)} = O\left\{ |u|^{-1} e^{\frac{u^2}{2}} \sum_0^\infty (m+1)^{-(f+1)} \right\}$$

$$= O\left(|u|^{-1} \; e^{\frac{u^2}{2}} \right)$$

et de même pour $\tau_n^{(\delta)}$:

$$\tau_n^{(\delta)} = \sum_0^n A_m^{-(f+1)}\; \tau_{n-m}^{(k)} = O\left\{ e^{\frac{u^2}{2}} \sum_0^\infty (m+1)^{-(f+1)} \right\} = O\left(e^{\frac{u^2}{2}} \right).$$

Les inégalités (AH) et (BH) sont d'autre part faciles à démontrer pour $|u| \leqq a$, quelque grand que soit le nombre fixe a. A cet effet nous utilisons le résultat de M. Maitland Wright[9] qui a donné la formule approchée du coefficient c_n, où

$$\sum_0^\infty c_n z^n = (1-z)^{-2l} \quad e^{\frac{\sigma\, e^{i\theta}}{1-z}} \quad f(z)\,,$$

$f(z)$ étant holomorphe pour $|z| \leqq 1 + \epsilon$, $\epsilon > 0$.

Cette formule, valable pour σ réel, positif et *fini*, s'écrit ainsi (Théorème II):

$$c_n = \frac{f(1)}{2\sqrt{\pi}}\, e^{\frac{\sigma}{2} e^{i\theta}} \quad (\sigma e^{i\theta})^{\frac{1}{4}-l} \quad n^{l-\frac{3}{4}}$$

$$\times \left\{ e^{2\sigma^{\frac{1}{2}} e^{i\frac{\theta}{2}}} \sqrt{\sigma n} + e^{-2\sigma^{\frac{1}{2}} e^{i\frac{\theta}{2}}} \sqrt{\sigma n}\, + (2l-\tfrac{1}{2})\pi i + O\left(\frac{1}{\sqrt{n}} \right) \right\}.$$

Les fonctions génératrices des suites $\{\sigma_n^{(\delta)}\}$ et $\{\tau_n^{(\delta)}\}$ sont du type général

$$G_\beta^{(\alpha)}(z) = \sum_0^\infty g_n z^n = (1-z)^{-2\alpha}\; e^{-\frac{d^2 z}{1-z}}\; (1+z)^{-2\beta}\; e^{\frac{z^2 z}{1+z}}$$

avec $\beta = \frac{1}{4}$ pour $\{\sigma_n^{(\delta)}\}$ et $\beta = \frac{3}{4}$ pour $\{\tau_n^{(\delta)}\}$. Plus exactement on a

$$\sqrt{\pi} \sum_0^\infty \sigma_n^{(\delta)} z^n = G_{\frac{1}{4}}^{(\alpha)}(z) \qquad \left(\alpha = \frac{\delta}{2} + \frac{3}{4} \right)$$

[9] *Journal of Lond. Math. Society*, vol. 7, 1932, pp. 261-262.

tandis que

$$\sqrt{\pi}\sum_0^\infty z^n\,\tau_{n+1}^{(\delta)} = (u-x)\,G_{\frac{1}{4}}^{\left(\alpha+\frac{1}{2}\right)}(z) + (u+x)\,G_{\frac{3}{4}}^{(\alpha)}(z)\,.$$

Dans les deux cas on a posé $\delta = 2\alpha - \frac{1}{2}$. En modifiant légèrement l'analyse de M. Maitland Wright on constate que sa formule reste applicable aussi à $g_n = g_n(\alpha, \beta)$ seulement on doit tenir compte des deux points singuliers $z_1 = +1$ et $z_2 = -1$ ce qui fait apparaître deux groupes de termes. Pour le point $z_1 = +1$ on a: $\sigma = d^2$, $\theta = \pi$, $l = \alpha$ et

$$f(z) = (1+z)^{-2\beta}\; e^{d^2+\frac{s^2\cdot z}{1+z}}$$

tandis que sur le point $z_2 = -1$ on pose $z = -\zeta$ et alors $\sigma = s^2$, $\theta = \pi$, $l = \beta$ et

$$f(\zeta) = (1+\zeta)^{-2\alpha}\; e^{s^2+\frac{d^2\zeta}{1+\zeta}}\,.$$

On obtient ainsi sous l'hypothèse que x et u sont *finis* tous les deux et tenant compte du fait que $s^2 + d^2 = u^2 + x^2$

$$e^{-\frac{s^2+u^2}{2}}\,g_n(\alpha,\beta)\sqrt{\pi} = d^{\frac{1}{2}-2\alpha}\,n^{\alpha-\frac{3}{4}}\left\{4^{-\beta}\cos\left[2d\sqrt{n} - (\alpha-\tfrac{1}{4})\,\pi\right] + O\left(n^{-\frac{1}{2}}\right)\right\}$$

$$+ (-1)^n\,s^{\frac{1}{2}-2\beta}\,n^{\beta-\frac{3}{4}}\left\{4^{-\alpha}\cos\left[2s\sqrt{n} - (\beta-\tfrac{1}{4})\,\pi\right] + O\left(n^{-\frac{1}{2}}\right)\right\}$$

d'où immédiatement l'inégalité

$$g_n = g_n(\alpha,\beta) = O\left\{\frac{e^{\frac{s^2+u^2}{2}}}{\sqrt{n}}\left[\left(\frac{\sqrt{n}}{d}\right)^{2\alpha-\frac{1}{2}} + \left(\frac{\sqrt{n}}{s}\right)^{2\beta-\frac{1}{2}}\right]\right\}.$$

En y posant $\beta = \dfrac{1}{4}$ on trouve, vu que $\alpha = \dfrac{\delta}{2} + \dfrac{3}{4}$:

$$H_n^{(\delta)}(x,u) = \frac{\sigma_n^{(\delta)}}{A_n^{(\delta)}} = \frac{\Gamma(\delta+1)\,g_n(\alpha,\frac{1}{4})}{n^\delta\sqrt{\pi}}\,[1+o(1)] = O\left[\frac{e^{\frac{s^2+u^2}{2}}}{d^{\delta+1}\sqrt{n^\delta}}\right]$$

car u et x étant finis, on a $d < \sqrt{n}$. De même, on vérifie facilement que

$$\frac{\partial H_n^{(\delta)}}{\partial x} = \frac{\tau_n^{(\delta)}}{A_n^{(\delta)}} = O\left[\frac{e^{\frac{s^2+u^2}{2}}\sqrt{n}}{d^{\delta+1}\sqrt{n^\delta}}\right].$$

Ces résultats prouvent (AH) et (BH) pour $|u| \leqq a$.

§4. Il nous reste maintenant à prouver (AH) et (BH) dans les intervalles $a \leqq |u| \leqq \eta \sqrt{n}$, le nombre a étant aussi grand qu'on veut, mais fixe. C'est ici qu'il est indispensable de recourir à la méthode du col. Nous allons l'appliquer à l'étude du coefficient g_n du développement de la fonction

$$G_\beta^{(\alpha)}(z) = \sum_0^\infty g_n z^n = \frac{e^{-\frac{d^2 z}{1-z} + \frac{s^2 z}{1+z}}}{(1-z)^{2\alpha}\ (1+z)^{2\beta}}$$

où l'on suppose α et β réels et quelconques. Quant à l'intervalle de variation de u, qu'on peut supposer positif sans nuire à la généralité, $a \leqq u \leqq \eta \sqrt{n}$, nous choisissons d'abord un nombre positif $\epsilon > 0$ aussi petit qu'on veut mais fixe et supposons ensuite a suffisamment grand et η suffisamment petit pour avoir

$$(1 - \sqrt{\epsilon})s^2 < d^2 < s^2 < A\,\epsilon \quad .$$

en posant pour abréger $A = n + 1 + 2\alpha + 2\beta$. Ce choix de l'intervalle est réalisé p. ex. si l'on prend $a\sqrt{\epsilon} > 4|x|$ et $\eta = \sqrt{\epsilon}$. Il entraîne évidemment $n > 8\epsilon^{-2}x^2$ (ce qui est toujours réalisé, car $n \to \infty$, tandis que x et ϵ sont fixes), ainsi que $d^{-2} = O(\epsilon)$ et $A^{-1} = O(\epsilon^2)$. Ceci posé on a

$$g_n = \frac{1}{2\pi i} \int_{(C)} F(z)\, dz$$

où

$$F(z) = z^{-(n+1)}\ (1-z)^{-2\alpha}\ (1+z)^{-2\beta}\ e^{-\frac{d^2 z}{1-z} + \frac{s^2 z}{1+z}}$$

le chemin d'intégration (C) entourant l'origine $z = 0$. Pour trouver les cols formons

$$\frac{F'(z)}{F(z)} = \frac{2\alpha}{1-z} - \frac{2\beta}{1+z} - \frac{n+1}{z} - \frac{d^2}{(1-z)^2} + \frac{s^2}{(1+z)^2}\,.$$

L'équation $F'(z) = 0$, donnant les cols, s'écrit:

$$(8)\begin{cases} \Phi(z) = Az^4 + (2\alpha - 2\beta + d^2 - s^2)z^3 \\ \qquad - 2[A - (d^2 + s^2 + \alpha + \beta)]z^2 - (2\alpha - 2\beta - d^2 + s^2)z \\ \qquad + A - 2(\alpha + \beta) = 0\,. \end{cases}$$

Pour résoudre cette équation en $z = \rho e^{i\theta}$ divisons le premier membre par $e^{2i\theta}$ et écrivons les deux équations réelles en ρ et θ:

$$(9)\quad \begin{aligned}&[A(1+\rho^4) - 2(\alpha+\beta)]\cos 2\theta - 2[A - (d^2+s^2+\alpha+\beta)]\rho^2\\ &+ [(2\alpha - 2\beta + d^2 - s^2)\rho^2 - (2\alpha - 2\beta - d^2 + s^2)]\rho\cos\theta = 0,\end{aligned}$$

$$(10)\quad \begin{aligned}&[A(1-\rho^4) - 2(\alpha+\beta)]\sin 2\theta\\ &= [(2\alpha - 2\beta + d^2 - s^2)\rho^2 - (2\alpha - 2\beta - d^2 + s^2)]\,\rho\sin\theta.\end{aligned}$$

On a par hypothèse $s^2 < A\epsilon$, la quantité positive ϵ étant fixe, mais aussi petite qu'on veut. Dans ces conditions l'équation (8) n'a pas de racines réelles, car les valeurs $z = \pm 1$ ne la vérifient point et pour $z^2 \neq 1$ on a

$$\Phi(z) = A(z^2-1)^2\,[1 + \vartheta(z)]$$

où $|\vartheta(z)| < 1$ pourvu que A soit suffisamment grand. Par conséquent, $\sin\theta \neq 0$ et de (10) on tire

$$(11)\quad 2\cos\theta = \rho\,\frac{s^2 - d^2 + 2\alpha - 2\beta - (s^2 - d^2 - 2\alpha + 2\beta)\,\rho^2}{A(1-\rho^4) - (2\alpha + 2\beta)}$$

ce qui permet d'écrire (9), en posant $\rho^2 = r$, ainsi:

$$(12)\quad r^6 + 2a_1 r^5 - a_2 r^4 - 2a_3 r^3 - a_4 r^2 + 2a_5 r + a_6 = 0$$

avec
$$a_1 = 1 + \frac{s^2 + d^2 + \gamma}{A}, \qquad a_5 = a_1\left(1 - \frac{2\gamma}{A}\right)^2$$

$$a_2 = 1 - \frac{2\gamma}{A} - \frac{(s^2 - d^2)^2 - 4\delta^2}{A^2}, \qquad a_4 = a_2\left(1 - \frac{2\gamma}{A}\right)$$

$$a_3 = 2 - 2\frac{s^2 + d^2 + 3\gamma}{A} + \frac{(s^2 - d^2)^2 + 4\gamma(s^2 + d^2) + 4(\gamma^2 + \delta^2)}{A^2} - \gamma\,\frac{(s^2 - d^2 - 2\delta)^2}{A^3}$$

$$a_6 = \left(1 - \frac{2\gamma}{A}\right)^3$$

où l'on a posé pour abréger $\alpha + \beta = \gamma$ et $\alpha - \beta = \delta$.

On constate que l'on a

$$\frac{a_4}{a_2} = \sqrt{\frac{a_5}{a_1}} = \sqrt[3]{\frac{a_6}{1}} = 1 - \frac{2\gamma}{A}$$

ce qui permet de réduire l'équation (12) à une équation réciproque de troisième degré et de la résoudre. On trouve ainsi, étant donné que $s^2 - d^2 = 2ux = O(u) = O(s) = O(\sqrt{A\epsilon})$, ainsi que $s^2 + d^2 = O(s^2) = O(A\epsilon)$:

$$\rho_1 = 1 - \frac{\alpha}{A} - \frac{\alpha s^2 - \beta d^2}{4A^2} + O(\epsilon^4)\,,$$

$$\rho_2 = 1 - \frac{\beta}{A} + \frac{\alpha s^2 - \beta d^2}{4A^2} + O(\epsilon^4)\,.$$

A ces valeurs de ρ correspondent les valeurs $\pm\theta_1$ et $\pm\theta_2$ de θ, où

$$\theta_1 = \frac{d}{\sqrt{A}}[1 + O(\epsilon)]\,, \qquad \theta_2 = \pi - \frac{s}{\sqrt{A}}[1 + O(\epsilon)]$$

et les quatre cols $\rho_1 e^{\pm i\theta_1}$, $\rho_2 e^{\pm i\theta_2}$ sont déterminés. On trouve également pour leurs coordonnées x_k, $\pm y_k$:

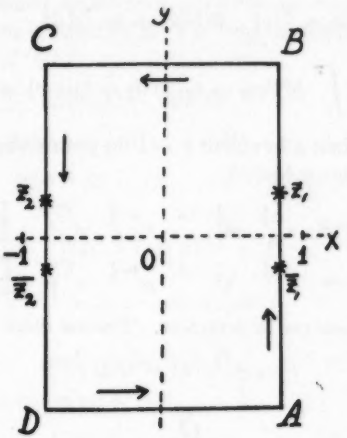
$$\begin{cases} x_1 = \rho_1 \cos\theta_1 = 1 - \dfrac{d^2}{2A} - \dfrac{s^2 d^2}{8A^2} + O(\epsilon^3) \\ y_1 = \rho_1 \sin\theta_1 = \dfrac{d}{\sqrt{A}}\left\{1 - \dfrac{\alpha - \beta}{2A} - \dfrac{\alpha}{d^2} + O(\epsilon^3)\right\} \end{cases}$$

$$\begin{cases} x_2 = \rho_2 \cos\theta_2 = -1 + \dfrac{s^2}{2A} + \dfrac{s^2 d^2}{8A^2} + O(\epsilon^3) \\ y_2 = \rho_2 \sin\theta_2 = \dfrac{s}{\sqrt{A}}\left\{1 + \dfrac{\alpha - \beta}{2A} - \dfrac{\beta}{s^2} + O(\epsilon^3)\right\}. \end{cases}$$

Ceci posé, on étudie la disposition des vallées de ces cols pour choisir le chemin d'intégration (C). Pour abréger nous allons voir qu'il suffit de choisir pour (C) les deux droites parallèles à l'axe OY et passant par les cols conjugués. Soient $z = x_1 + iy_1\xi$ et $z = x_2 + iy_2\xi$, ξ variant de $-\infty$ à $+\infty$ et posons

$$F(z) = G(\xi)\, e^{iH(\xi)}\,.$$

Pour prouver qu'on peut réduire (C) à deux droites parallèles à OY et passant par les cols considérons le rectangle $ABCD$, où $\overline{AB} = \overline{CD} = 2T = 2d^{-\frac{1}{2}}\sqrt[4]{A}$.



Sur les segments $\overline{BC}$ et $\overline{DA}$ on a $z = w \pm i\,T$, w variant de x_2 à x_1, donc

$$i_1 = \frac{1}{2\pi i}\int_{\overline{BC}} F(z)dz = -\frac{1}{2\pi i}\int_{x_2}^{x_1} F(w + iT)\,dw$$

et

$$i_2 = \frac{1}{2\pi i}\int_{\overline{DA}} F(z)dz = \frac{1}{2\pi i}\int_{x_2}^{x_1} F(w - iT)\,dw$$

de sorte que

$$\pi\,|\,i_1 + i_2\,| = \left|R\left(i\int_{x_2}^{x_1} F(w + iT)\,dw\right)\right| \leqq \int_{x_2}^{x_1} |\,F(w + iT)\,|\,dw.$$

Or, pour $|\,w\,| \leqq 1$

$$|\,F(w + iT)\,| < \frac{e^{s^2+d^2}}{T^A}$$

et par conséquent, vu que $s^2 + d^2 < 2A\epsilon$, $T^{-1} = d^3\,A^{-\frac{1}{4}}$ donc $T^{-1} < \sqrt[4]{\epsilon}$:

$$\pi\,|\,i_1 + i_2\,| < 2\,e^{-\frac{A}{4}\left[\log\frac{1}{\epsilon} - 8\epsilon\right]}.$$

On peut toujours supposer ϵ assez petit pour avoir $-\log\epsilon - 8\epsilon > 4$ et ainsi

$$g_n = \frac{1}{2\pi i}\int_{(c)} F(z)dz = \frac{y_1}{\pi}\int_0^{\tau} R[F(x_1 + iy_1\,\xi)]\,d\xi$$

$$-\frac{y_2}{\pi}\int_0^{\tau} R[F(x_2 + iy_2\,\xi)\,d\xi + O(e^{-n}) = J_1 + J_2 + O(e^{-n})$$

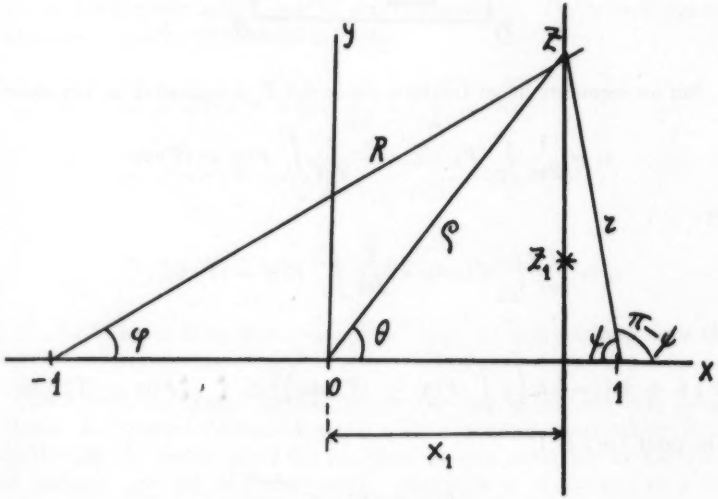
les cols correspondant à la valeur $\xi = 1$ du paramètre ξ. Les intégrales J_1 et J_2 vérifient les inégalités

$$(13)\qquad |J_1| < 2^{1-2\beta}\ \pi^{-\frac{1}{2}}\ d^{\frac{1}{2}-2\alpha}\ n^{\alpha-\frac{3}{4}}\ e^{\frac{d^2+s^2}{2}}\left[1 + O\left(\epsilon^{\frac{3}{4}}\right)\right]$$

$$(14)\qquad |J_2| < 2^{1-2\alpha}\ \pi^{-\frac{1}{2}}\ s^{\frac{1}{2}-2\beta}\ n^{\beta-\frac{3}{4}}\ e^{\frac{d^2+s^2}{2}}\left[1 + O\left(\epsilon^{\frac{3}{4}}\right)\right]$$

et il suffit de démontrer la première. Posons donc

$$F(x_1 + iy_1\,\xi) = G(\xi)\,e^{iH(\xi)}$$



et soient $z + 1 = Re^{i\varphi}$, $z = \rho e^{i\theta}$, $z = 1 + re^{i(\pi-\psi)}$, donc

$$G(\xi) = R^{-2\beta}\ r^{-2\alpha}\ \rho^{-(n+1)}\ e^{-d^2\ r^{-2}(x_1-\rho^2)+s^2\ R^{-2}\ (x_1+\rho^2)},$$

$$H(\xi) = (R^{-2}\,s^2 - r^{-2}\,d^2)\,y_1\,\xi - 2\alpha(\pi - \psi) - 2\beta\varphi - (n+1)\,\theta$$

d'où $G(-\xi) = G(\xi)$, $H(-\xi) = -H(\xi)$ et par conséquent

$$J_1 = \frac{y_1}{\pi}\int_0^{\tau} G(\xi)\cos[H(\xi)]\,d\xi\,.$$

Il importe de vérifier avant tout que $G'(\xi)$ ne s'annule qu'une seule fois, pour $\xi = 1$, dans l'intervalle $(0, T)$. On trouve facilement l'expression suivante pour $G'(\xi)$:

$$R^4 r^4 \rho^2 y_1^{-2} \xi^{-1} G^{-1}(\xi) G'(\xi) = 16\Psi(t)$$
$$= 2d^2(1 - x_1) R^4 \rho^2 + 2s^2(1 + x_1) r^4 \rho^2 - 2\alpha\rho^2 r^2 R^4$$
$$- 2\beta\rho^2 r^4 R^2 - (A - 2\alpha - 2\beta) R^4 r^4$$

où nous avons posé

$$2tx_1 = 1 + x_1^2 + y_1^2 \xi^2$$

ce qui donne $t \geqq \dfrac{1 + x_1^2}{2x_1} = t_0 > 1$ ainsi que

$$\rho^2 = 2tx_1 - 1, \qquad R^2 = 2x_1(t + 1), \qquad r^2 = 2x_1(t - 1).$$

Le col $\xi = \xi_1 = 1$ correspond à

$$t = t_1 = \frac{1 + x_1^2 + y_1^2}{2x_1} = 1 + \frac{d^2}{2A} + O(\epsilon^2).$$

Avec cette nouvelle variable t, on trouve pour $\Psi(t)$ l'expression suivante

$$\Psi(t) = - A x_1^4 \psi(t) = - A x_1^4 [t^4 + C_1 t^3 + C_2 t^2 + C_3 t + C_4],$$

où

$$C_1 = - \frac{2s^2}{A} + O(\epsilon^2),$$

$$C_2 = - 2\left[1 - \frac{5s^2}{2A} + O(\epsilon^2)\right],$$

$$C_3 = - \frac{4s^2}{A} + O(\epsilon^2),$$

$$C_4 = 1 + \frac{s^2}{A} + O(\epsilon_2).$$

On a évidemment $\psi(t_1) = 0$, car $G'(\xi_1) = G'(1) = 0$. Ensuite, on calcule, étant donné la valeur de t_1, que

$$\psi'(t_1) = \frac{4d^2}{A} + O(\epsilon^2),$$

$$\psi''(t_1) = 8 + \frac{6d^2 - 2s^2}{A} + O(\epsilon^2),$$

$$\psi'''(t_1) = 24 + 12\frac{d^2 - s^2}{A} + O(\epsilon^2),$$

d'où enfin, en posant encore $\tau = t - t_1$

$$\psi(t) = \tau\chi(\tau) = \tau\left\{\tau^3 + [4 + O(\epsilon)]\tau^2 + [4 + O(\epsilon)]\tau + \frac{4d^2}{A} + \cdots\right\}$$

ainsi que

$$\chi'(\tau) = [3 + O(\epsilon)]\,\tau^2 + [8 + O(\epsilon)]\,\tau + 4 + O(\epsilon)\,.$$

On constate, par conséquent, que l'équation $\tau\chi(\tau) = 0$ n'a, sauf la racine évidente $\tau_1 = 0$, qu'une seule racine réelle τ^* et que

$$\tau^* = -\frac{d^2}{A} + O(\epsilon^2)\,.$$

La valeur correspondante de t est égale à

$$t^* = t_1 + \tau^* = 1 - \frac{d^2}{2A} + O(\epsilon^2)$$

et l'on voit que $t^* < 1$. Or, la variable t varie entre $t_0 > 1$ et t_2, où $t_2 = 1 + \dfrac{d}{2\sqrt{A}} + \cdots$. Donc dans l'intervalle (t_0, t_2) la fonction $\Psi(t)$ ne s'annule qu'une seule fois, au col. On en déduit que $G'(\xi)$ s'annule dans l'intervalle $(0, T)$: pour $\xi = 0$ et pour $\xi = \xi_1 = 1$, c'est à dire au col, et la fonction $G(\xi)$ croît quand ξ va de 0 à 1, atteint son maximum au col et décroit ensuite constamment.

Etudions de même la dérivée $H'(\xi)$, dont l'expression à l'aide de la variable t est la suivante:

$$-x_1^2(t^2 - 1)^2(2tx_1 - 1)\,H'(\xi) = Ay_1 f(t)$$
$$= Ay_1\{c_0t^4 + c_1t^3 + c_2t^2 + c_3t + c_4\}\,,$$

où

$$c_0 = 1 + \frac{2s^2 - 5d^2}{2A} + O(\epsilon^2)$$

$$c_1 = -\frac{11s^2 + d^2}{2A} + O(\epsilon^2)$$

$$c_2 = -2\left[1 - \frac{19s^2 + 9d^2}{4A} + O(\epsilon^2)\right]$$

$$c_3 = -\frac{13s^2 - d^2}{2A} + O(\epsilon^2)$$

$$c_4 = 1 + \frac{3s^2 - 4d^2}{2A} + O(\epsilon^2)\,.$$

On a évidemment $f(t_1) = 0$, car $H'(\xi)$ s'annule au col. Posons à nouveau $\tau = t - t_1$ et calculons les dérivées:

$$f'(t_1) = \frac{2d^2}{A} + O(\epsilon^2)$$

$$f''(t_1) = 8 - 2\frac{6d^2 + s^2}{A} + O(\epsilon^2)$$

$$f'''(t_1) = 24 - 3\frac{3s^2 + 17d^2}{A} + O(\epsilon^2)\,.$$

Par conséquent, on trouve

$$f(t) = \tau\varphi(\tau) = \tau\Big\{[1 + O(\epsilon)]\tau^3 + [4 + O(\epsilon)]\tau^2 + [4 + O(\epsilon)]\tau + \frac{2d^2}{A} + O(\epsilon^2)\Big\}$$

et

$$\varphi'(\tau) = [3 + O(\epsilon)]\tau^2 + [8 + O(\epsilon)]\tau + 4 + O(\epsilon)\,.$$

On constate que l'équation $\varphi(\tau) = 0$ n'a qu'une racine réelle qui est égale à τ_*, où

$$\tau_* = -\frac{d^2}{2A} + O(\epsilon^2)\,.$$

Cela veut dire que $H'(\xi)$ s'annule pour $\xi = \xi_*$, où

$$y_1\,\xi_* = \sqrt{2x_1\tau_* - (1 - 2x_1t_1 + x_1^2)}\,.$$

En poussant les calculs approchés plus loin on peut à partir de cette expression de ξ_* établir que l'on a

$$\xi_* = \frac{d}{2\sqrt{A}}[1 + O(\epsilon)]\,,$$

mais pour abréger nous allons vérifier directement que

$$\xi_* < \frac{1 - x_1}{y_1}\left(1 + \frac{d^2}{A}\right) = \xi_2. \tag{15}$$

Le signe de $H'(\xi)$ est opposé à celui de $\tau\varphi(\tau)$, donc on a $H'(\xi) < 0$ pour $\xi > 1$ et $H'(\xi) > 0$ pour $\xi_* < \xi < 1$. Or, on va voir que $H'(\xi)$

est déjà positive pour $\xi = \xi_2$ ce qui suffit pour prouver (15). Dans l'expression de $H'(\xi)$

$$H'(\xi) = -\frac{(n+1)x_1}{\rho^2} + \frac{d^2}{r^4}[y_1^2\xi^2 - (1-x_1)^2] + \frac{2\alpha(1-x_1)}{r^2} + \frac{s^2}{R^2}\left[1 - \frac{2y_1^2\xi^2}{R^2} - \frac{2\beta(1+x_1)}{s^2}\right].$$

Le troisième et le quatrième termes sont positifs car $s^{-2} = O(\epsilon)$. En les omettant, en posant

$$\xi = \frac{1-x_1}{y_1}\left(1 + \lambda\frac{d^2}{A}\right)$$

d'où

$$r^2 = (1-x_1)^2 + y_1^2\xi^2 = 2(1-x_1)^2\left\{1 + \frac{\lambda d^2}{A} + \frac{\lambda^2 d^4}{2A^2}\right\}$$

et en observant que $\rho \geqq x_1$, on trouve

$$H'(\xi) > -\frac{A}{x_1} + \frac{d^2}{4(1-x_1)^2}\,\frac{\lambda d^2}{A}\left(2 + \frac{\lambda d^2}{A}\right)\left\{1 + \frac{\lambda d^2}{A} + \frac{\lambda^2 d^4}{2A^2}\right\}^{-2}.$$

A l'aide de l'expression approchée pour x_1 on en déduit que pour $2\lambda > 1$ on a

$$H'(\xi) > (2\lambda - 1)\,A\,[1 + O(\epsilon)] > 0$$

ce qui prouve notre assertion (15). Vu que

$$\frac{1-x_1}{y_1} = \frac{d}{2\sqrt{A}}[1 + O(\epsilon)]$$

on en déduit que $\xi^* < \dfrac{d}{2\sqrt{A}}[1 + O(\epsilon)]$. En effet, on établit facilement la valeur de $H'(0)$:

$$H'(0) = \frac{4A^2}{d^2}[-1 + O(\epsilon)] < 0$$

ce qui prouve qu'effectivement $0 < y_1\xi^* < (1-x_1)\,[1 + O(\epsilon)]$.

Maintenant nous pouvons conclure:

$$\pi J_1 = y_1\int_0^{\frac{d}{\sqrt{A}}} + y_1\int_{\frac{d}{\sqrt{A}}}^{\frac{\sqrt[4]{A}}{\sqrt{d}}} = i_1 + i_2\,.$$

On a

$$|i_1| < y_1 G\left(\frac{d}{\sqrt{A}}\right)\frac{d}{\sqrt{A}} = \frac{d^2}{A}G\left(\frac{d}{\sqrt{A}}\right)[1+O(\epsilon)].$$

Or,

$$G\left(\frac{d}{\sqrt{A}}\right) = O[e^{\frac{d^2+s^2}{2}}\ A^{2\alpha}\ d^{-4\alpha}\ e^{-\frac{2A}{5}(1+o(\epsilon))}],$$

ce qui nous donne, vu que $A \sim n$:

$$|i_1| < e^{\frac{d^2+s^2}{2}}\ d^{\frac{1}{2}-2\alpha}\ n^{\alpha-\frac{3}{4}}\ O\left(n^{\alpha-\frac{1}{4}}d^{\frac{3}{2}-2\alpha}\ e^{-\frac{2n}{5}}\right)$$

c'est à dire pour $n \to \infty$

$$|i_1| = e^{\frac{d^2+s^2}{2}}\ d^{\frac{1}{2}-2\alpha}\ n^{\alpha-\frac{3}{4}}\ o\left(e^{-\frac{n}{3}}\right).$$

Quant à l'intégrale

$$i_2 = y_1\int_{\frac{d}{\sqrt{A}}}^{\sqrt[4]{\frac{A}{d}}} G(\xi)\ \cos[H(\xi)]\ d\xi = y_1\int_{\frac{d}{\sqrt{A}}}^{1} + y_1\int_{1}^{\sqrt[4]{\frac{A}{d}}} = i_3 + i_4$$

nous allons l'évaluer en appliquant aux intégrales i_3 et i_4 le lemme du § 1. A cet effet, observons que pour $\xi \leqq 1$ dans l'expression de $H''(\xi)$

$$\frac{H''(\xi)}{2y_1^3\xi} = \frac{n+1}{\rho^4} + \frac{s^2+2\beta(1+x_1)}{R^4} + \frac{4d^2(1-x_1)^2}{r^6} - \frac{4s^2(1+x_1)^2}{R^6} - \frac{d^2+2\alpha(1-x_1)}{r^4}$$

les deux premiers termes sont respectivement $O(A)$ et $O(s^2)$. Le quatrième terme est aussi $O(s^2) = O(A\epsilon)$. Ensuite

$$r^{-4}[d^2+2\alpha(1-x_1)] = r^{-4}d^2\left(1+\frac{\alpha}{A}[1+O(\epsilon)]\right) = r^{-4}d^2[1+O(\epsilon^2)]$$

tandis que

$$4r^{-6}\,d^2\,(1-x_1)^2 = r^{-6}d^6\,A^{-2}[1+O(\epsilon)].$$

Par conséquent, le troisième et le dernier termes donnent

$$-d^2r^{-6}\{r^2[1+O(\epsilon^2)] - d^4A^{-2}[1+O(\epsilon)]\},$$

où

$$r^2 = (1 - x_1)^2 + y_1^2 \xi^2 = \tfrac{1}{4} d^4 A^{-2} [1 + O(\epsilon)] + d^2 \xi^2 A^{-1} [1 + O(\epsilon)].$$

Ainsi, on a trouvé que

$$\frac{H''(\xi)}{2y_1^3 \xi} = -\frac{d^4}{A r^6} \left(\xi^2 - \tfrac{3}{4} \frac{d^2}{A}\right) [1 + O(\epsilon)].$$

Or, dans i_3 on a $\xi \geqq \dfrac{d}{\sqrt{A}}$, d'où enfin

$$H''(\xi) < 0 \qquad \left(\frac{d}{\sqrt{A}} \leqq \xi \leqq 1\right).$$

En outre pour $\xi = 1$ on a $r \sim y_1$. Par conséquent

$$H''(1) = \frac{2d^4 A \sqrt{A}}{A d^3} [1 + O(\epsilon)] = 2d \sqrt{A} [1 + O(\epsilon)].$$

Calculons ensuite $H'''(\xi)$. On trouve, tenant compte du fait que $rr' = \rho\rho' = RR' = y_1^2 \xi$:

$$\begin{aligned} 2^{-1} y_1^{-3} H'''(\xi) &= (n + 1)(x_1^2 - 3y_1^2 \xi^2) \rho^{-6} \\ &+ [s^2 + 2\beta(1 + x_1)][(1 + x_1)^2 - 3y_1^2 \xi^2] R^{-6} \\ &+ 4d^2 (1 - x_1)^2 [(1 - x_1)^2 - 5y_1^2 \xi^2] r^{-8} \\ &- 4s^2 (1 + x_1)^2 [(1 + x)^2 - 5y_1^2 \xi^2] R^{-8} \\ &- [d^2 + 2\alpha (1 - x_1)][(1 - x_1)^2 - 3y_1^2 \xi^2] r^{-6}. \end{aligned}$$

Soit $\xi_0 \leqq \xi < \dfrac{1}{d} \sqrt{\dfrac{A}{3}}$ où, $\xi_0 > 0$ est une quantité fixe inférieure à l'unité. On trouve

$$\frac{H'''(\xi)}{2y_1^3} > \frac{3 d^4 \xi^2}{A r^8} \left(r^2 - \tfrac{5}{3} \frac{d^4}{A^2}\right) [1 + O(\epsilon)] - \frac{s^2}{16} [1 + O(\epsilon)].$$

Or,

$$r^2 - \tfrac{5}{3} \frac{d^4}{A^2} = \frac{d^2 \xi^2}{A} [1 + O(\epsilon)] \quad \text{et} \quad r^2 \geqq \frac{d^2 \xi^2}{A} [1 + O(\epsilon)].$$

Par conséquent

$$\frac{H'''(\epsilon)}{2y_1^3} > \left[27d^2 - \frac{s^2}{16}\right][1 + O(\epsilon)] > 0\,,$$

si ξ est compris dans l'intervalle $\left(\xi_0, \frac{1}{d}\sqrt{\frac{A}{3}}\right)$.

Ceci posé, on constate que le lemme est applicable à l'intégrale i_3 et l'on trouve ainsi, vu que $H'(1) = 0$:

$$|\,i_3\,| < y_1\,G(1)\sqrt{\frac{2\pi}{|\,H''(1)\,|}} = \frac{d\,G(1)}{\sqrt{A}}\,\frac{\sqrt{\pi} + O(\epsilon)}{\sqrt{d\sqrt{A}}}\,,$$

où

$$G(1) = 4^{-\beta}\,e^{\frac{d^2+s^2}{2}}\,d^{-2\alpha}\,A^{\alpha}\,[1 + O(\epsilon)]\,.$$

Par conséquent

$$|\,i_3\,| < 4^{-\beta}\,\sqrt{\pi}\;e^{\frac{d^2+s^2}{2}}\;d^{\frac{1}{2}-2\alpha}\;A^{\alpha-\frac{3}{4}}\;[1 + O(\epsilon)]\,.$$

Passons à l'intervalle $1 \leqq \xi \leqq T = d^{-\frac{1}{2}}\,A^{\frac{1}{4}}$. Dans cet intervalle on a constamment $H'''(\xi) > 0$ donc $H''(\xi)$ croît. Pour pouvoir appliquer le lemme à l'intégrale i_4 il faut d'abord démontrer que $H''(\xi)$ garde son signe constant, c'est à dire reste négative dans $(1, T)$. Pour cela il suffit de montrer que $H''(T)$ est encore négative. Pour $\xi = T$ on trouve $y_1^2\,\xi^2 = y_1^2 T^2 = dA^{-\frac{1}{2}}\left[1 + O\left(\epsilon^{\frac{3}{2}}\right)\right]$ ce qui entraine

$$(n+1)\,\rho^{-4} = A\,[1 + O(\epsilon^2)]\;\rho^{-4} = A - 2d\sqrt{A} + O(d^2)\,.$$

Ensuite

$$4d^2\,(1 - x_1)^2\,r^{-6} + [s^2 + 2\beta(1 + x_1)]\,R^{-4} - 4s^2\,(1 + x_1)^2\,R^{-6} = O(d^2)$$

et

$$[d^2 + 2\alpha(1 - x_1)]\,r^{-4} = A + O(d^2\sqrt{\epsilon})\,.$$

Par conséquent

$$y_1^{-\frac{5}{2}}\,H''(T) = -4d\sqrt{A}\,[1 + O(\epsilon)] < 0\,.$$

Pour appliquer le lemme à i_4 posons $G(\xi) = \vartheta(\xi)$ et $f(\xi) = -H(\xi)$, d'où $\vartheta' < 0, f' > 0, f'' > 0$ et $f''' < 0$. Ainsi le lemme est applicable et l'on obtient pour i_4 la même inégalité que pour i_3.

Par conséquent l'inégalité (13) est démontrée. La preuve de (14) est exactement la même et nous l'omettons. On a trouvé ainsi pour le coefficient g_n de la série

$$\sum_0^\infty g_n(\alpha, \beta)\, z^n = (1-z)^{-2\alpha}\,(1+z)^{-2\beta}\, e^{-\frac{d^2 z}{1-z} + \frac{s^2 z}{1+z}}$$

l'inégalité

$$|g_n(\alpha,\beta)| < \frac{2}{\sqrt{\pi}}\, e^{\frac{d^2+s^2}{2}} \left\{ \frac{n^{\alpha-\frac{3}{4}}}{4^\beta d^{2\alpha-\frac{1}{2}}} + \frac{n^{\beta-\frac{3}{4}}}{4^\alpha s^{2\beta-\frac{1}{2}}} \right\} [1 + O(\epsilon)]$$

valable pour

$$(1-\sqrt{\epsilon})s^2 < d^2 < s^2 < \epsilon\, n\,.$$

Or, on a $\sqrt{\pi}\,\sigma_n^{(\delta)}(x,u) = g_n(\alpha, \frac{1}{4})$ ainsi que

$$\sqrt{\pi}\,\tau_n^{(\delta)}(x,u) = (u-x)g_n(\alpha+\tfrac{1}{2}, \tfrac{1}{4}) + (u+x)g_n(\alpha, \tfrac{3}{4})$$

ce qui prouve les inégalités fondamentales (AH) et (BH) pour tout u, x variant par hypothèse dans un intervalle *fini*, $|x| \leqq a$, et u extérieur à l'intervalle $(x - n^{-\frac{1}{2}}, x + n^{-\frac{1}{2}})$.

D'ailleurs on peut lever la dernière restriction a pour $|u-x|\sqrt{n} \leqq 1$, c'est à dire pour $(d\sqrt{n})^{-1} \geqq 1$:

$$\text{(AH)} \quad \begin{cases} |H_n^{(\delta)}(x,u)| \leqq \max\limits_{0 \leqq m \leqq n} |H_m^{(0)}(x,u)| = O\left(d^{-1}\, e^{\frac{x^2+u^2}{2}}\right) \\ \qquad = O\left\{d^{-1} e^{\frac{x^2+u^2}{2}} (d\sqrt{n})^{-\delta}\right\} = O\left[e^{\frac{x^2+u^2}{2}}\, n^{-\frac{\delta}{2}}\, d^{-(\delta+1)}\right] \end{cases}$$

ainsi que

$$\text{(BH)} \quad \begin{cases} \left|\frac{\partial}{\partial x} H_n^{(\delta)}(x,u)\right| \leqq \max\limits_{0 \leqq m \leqq n} \left|\frac{\partial}{\partial x} H_m^{(0)}(x,u)\right| = O\left(d^{-1}\, e^{\frac{x^2+u^2}{2}} \sqrt{n}\right) \\ \qquad = O\left[d^{-1} e^{\frac{x^2+u^2}{2}} \sqrt{n}\, (d\sqrt{n})^{-\delta}\right] = O\left[e^{\frac{x^2+u^2}{2}}\, n^{\frac{1-\delta}{2}}\, d^{-(\delta+1)}\right] \end{cases}$$

pourvu que l'on ait $\delta \geqq 0$.

§5. INÉGALITÉS FONDAMENTALES POUR LA MOYENNE $L_n^{(\delta)}(x, u)$ DE LA SÉRIE-NOYAU (L) ET POUR SA DÉRIVÉE PARTIELLE PAR RAPPORT À x

Nous commençons par déterminer la fonction génératrice $\phi_\alpha(z)$ de la série (L). On le fait le plus simplement à l'aide du calcul symbolique d'Heaviside. Soit $x \doteqdot p^{-1}$ alors

$$x^{n+\alpha}\, e^{-x} \doteqdot \Gamma(n+\alpha+1)\, p\, (p+1)^{-(n+\alpha+1)}$$

ce qui entraîne

$$n!\, e^{-x} x^\alpha L_n^{(\alpha)}(x) = \frac{d^n}{dx^n}(x^{n+\alpha}\, e^{-x}) \doteqdot \frac{p^{n+1}\Gamma(n+\alpha+1)}{(p+1)^{n+\alpha+1}}\,.$$

A l'aide de la relation bien connue

$$\sum_0^\infty w^n L_n^{(\alpha)}(u) = (1-w)^{-(\alpha+1)}\, e^{-\frac{uw}{1-w}}$$

on trouve:

$$x^\alpha e^{-x} \sum_0^\infty \frac{\Gamma(n+1)\, L_n^{(\alpha)}(x)\, L_n^{(\alpha)}(u)}{\Gamma(n+\alpha+1)}\, z^n = x^\alpha e^{-x}\, \phi_\alpha(z)$$

$$\doteqdot \frac{p}{(p+1)^{\alpha+1}} \sum_0^\infty \left(\frac{pz}{p+1}\right)^n L_n^{(\alpha)}(u) = \frac{pe^{-\frac{upz}{p+1-pz}}}{(p+1-pz)^{\alpha+1}}\,.$$

Posons $(1-z)^{-1} = a$ et $uz(1-z)^{-2} = b^2$. Calcul fait, on trouve

$$x^\alpha\, e^{-z}\phi_\alpha(z) \doteqdot \frac{e^{-\frac{uz}{1-z}}}{(1-z)^{\alpha+1}}\, \frac{pe^{\frac{b^2}{p+a}}}{(p+a)^{\alpha+1}}\,.$$

Or, de $(b^2x)^k \doteqdot k!\, p^{-k} b^{2k}$ on déduit

$$(b^{-2}x)^{\alpha/2}\, I_\alpha\,(2b\sqrt{x}) = (b^{-2}x)^{\alpha/2}\, J_\alpha(2bi\sqrt{x})$$

$$= b^{-2\alpha} \sum_0^\infty \frac{(b^2x)^{n+\alpha}}{n!\,(n+\alpha)!} \doteqdot b^{-2\alpha} \sum_0^\infty \frac{1}{n!}\left(\frac{b^2}{p}\right)^{n+\alpha} = p^{-\alpha}\, e^{\frac{b^2}{p}} = h(p)$$

d'où ensuite

$$\frac{p}{p+a}\, h(p+a) = \frac{pe^{\frac{b^2}{p+a}}}{(p+a)^{\alpha+1}} \doteqdot e^{-ax}\left(\frac{x}{b^2}\right)^{\alpha/2} I_\alpha\,(2b\sqrt{x})\,.$$

Par conséquent, tout calcul fait, on trouve:

$$x^{\alpha}\, e^{-x}\, \phi_{\alpha}(x) = \frac{e^{-\frac{u\,z}{1-z}}}{(1-z)^{\alpha+1}}\, e^{-\frac{x}{1-z}} \left(\frac{x(1-z)^2}{uz}\right)^{\alpha/2} I_{\alpha}\left(\frac{2\sqrt{uxz}}{1-z}\right)$$

c'est à dire

$$\phi_{\alpha}(z) = \frac{e^{-\frac{(u+x)\,z}{1-z}}}{(1-z)\sqrt{(uxz)^{\alpha}}}\, I_{\alpha}\left(\frac{2\sqrt{uxz}}{1-z}\right).$$

Désignons les sigma sommes de la série (L) par $\sigma_m^{(\delta)}(x, u)$ de sorte que

$$\sum_0^{\infty} z^n \sigma_n^{(\delta)}(x, u) = (1-z)^{-(\delta+1)}\, \phi_{\alpha}(z).$$

Le but de ce paragraphe sont les formules

$$\text{(AL)} \qquad L_n^{(\delta)}(x, u) = \frac{\sigma_n^{(\delta)}(x, u)}{A_n^{(\delta)}} = O\left\{\frac{e^{\frac{u+x}{2}}\,(ux)^{-\frac{\alpha}{2}-\frac{1}{4}}}{|\sqrt{u}-\sqrt{x}|^{\delta+1}\sqrt{n^{\delta}}}\right\}$$

et

$$\text{(BL)} \qquad \frac{\partial L_n^{(\delta)}(x, u)}{\partial x} = O\left\{\frac{e^{\frac{x+u}{2}}\,(ux)^{-\frac{\alpha}{2}-\frac{1}{4}}\sqrt{n}}{\sqrt{x}\,|\sqrt{u}-\sqrt{x}|^{\delta+1}\sqrt{n^{\delta}}}\right\}.$$

Soit d'abord $u \geqq e^2 n$. En développant la fonction de Bessel $I_{\alpha}[2(1-z)^{-1}\sqrt{uxz}]$, nous trouvons:

$$\sum_0^{\infty} \sigma_n^{(\delta)}(x, u) z^n = \sum_{n=0}^{\infty} \frac{(xu)^n}{n!\,\Gamma(n+\alpha+1)}\, \frac{e^{-\frac{(u+x)\,z}{1-z}}\, z^n}{(1-z)^{2n+\alpha+\delta+2}}.$$

Or,

$$\frac{e^{-\frac{(u+x)\,z}{1-z}}}{(1-z)^{2n+\alpha+\delta+2}} = \sum_{m=0}^{\infty} z^m L_m^{(2n+\alpha+\delta+1)}(x+u)$$

et par conséquent pour $x > 0$:

$$\sigma_n^{(\delta)}(x, u) = A_n^{(\delta)} L_n^{(\delta)}(x, u)$$

$$= \sum_{m=0}^{n} \frac{(xu)^m}{m!\,\Gamma(m+\alpha+1)}\, L_{n-m}^{(2m+\alpha+\delta+1)}(x+u)$$

tandis que par $x = 0$ on a

$$\Gamma(\alpha+1)\sigma_n^{(\delta)}(0, u) = L_n^{(\alpha+\delta+1)}(u) \qquad (x=0).$$

Le cas $x = 0$ est très simple et nous supposons dans la suite $x > 0$. Dans l'intervalle (e^2n, ∞) on a $x + u > u \geqq e^2n > 2n + \alpha + \delta + 1$ et cela entraine vu que $2(x + u) > 4n$ et grâce à (3) §2:

$$| L_{n-m}^{(2m+\alpha+\delta+1)}(x+u) | \leqq \frac{[2(x+u)]^{n-m}}{(n-m)!} < \frac{2(x+u)]^n}{4^m\, n!} \quad (0 \leqq m \leqq n)$$

d'où

$$\begin{aligned} | \sigma_n^{(\delta)}(x,u) | &< \frac{[2(x+u)]^n}{n!} \sum_{m=0}^{\infty} \frac{\left(\frac{xu}{4}\right)^m}{m!\,(m+\alpha)!} \\ &= \frac{2^{n+\alpha}(x+u)^n}{n!}\,(ux)^{-\frac{\alpha}{2}} I_\alpha(\sqrt{xu}) \\ &< \frac{2^{n+\alpha}(x+u)^n (ux)^{-\frac{\alpha}{2}} e^{\sqrt{xu}}}{n!\sqrt{2\pi\sqrt{xu}}} \left[1 + O\left(\frac{1}{\sqrt{xu}}\right)\right]. \end{aligned}$$

On trouve par conséquent vu que x est fini et $u \geqq e^2n$:

$$\begin{aligned} L_n^{(\delta)}(x,u) &= O\left[\, 2^n\, e^{n+\sqrt{xu}}\, u^{n-\frac{\alpha}{2}-\frac{1}{4}}\, n^{-\left(n+\delta+\frac{1}{2}\right)}\right] \\ &= O\left[u^{-\left(\frac{\alpha}{2}+\delta+\frac{3}{4}\right)}\, e^{\frac{u}{2}}\, \varphi_n(u)\right] \end{aligned}$$

où

$$\varphi_n(u) = 2^n e^{n+\sqrt{xu}-\frac{u}{2}} u^{n+\delta+\frac{1}{2}} n^{-\left(n+\delta+\frac{1}{2}\right)}.$$

Or, pour $u > 2n\left[1 + \sqrt{\frac{x}{n}} + O\left(\frac{1}{n}\right)\right]$ on a $\varphi_n'(u) < 0$ et d'autre part, comme on le vérifie facilement,

$$\varphi_n(e^2n) = e^{2n+1+2\delta}\, 2^n e^{n+e\sqrt{nx}-\frac{ne^2}{2}} < 1.$$

Par conséquent, la preuve de (AL) pour $u \geqq e^2n$ est faite. De même pour l'inégalité (BL) concernant l'expression

$$\frac{\partial \sigma_n^{(\delta)}(x,u)}{\partial x} = \sum_{m=0}^{n-1} \frac{(xu)^m}{m!} \cdot \left\{ \frac{u L_{n-m-1}^{(2m+\alpha+\delta+3)}(x+u)}{\Gamma(m+\alpha+2)} - \frac{L_{n-m-1}^{(2m+\alpha+\delta+2)}(x+u)}{\Gamma(m+\alpha+1)} \right\}$$

et dont nous omettons la preuve pour $u \geqq e^2n$, car elle ne diffère point de la preuve précédente.

§6

Dans ce paragraphe nous allons exposer la preuve des inégalités (AL) et (BL) dans l'intervalle $\eta n \leqq u < e^2 n$. Dans ce but nous allons d'abord étudier les moyennes d'ordre *entier* $\delta = [\delta] = k \geqq 0$.

Considérons les deux relations

$$L_{n+1}^{(\alpha)}(x) = L_n^{(\alpha)}(x) + L_{n+1}^{(\alpha-1)}(x),$$

$$xL_n^{(\alpha)}(x) = (n+\alpha)L_n^{(\alpha-1)}(x) - (n+1)\,L_{n+1}^{(\alpha-1)}(x)$$

et posons pour $k = [k] \geqq 0$ et $n = [n]$

$$G_n^{(\alpha,k)} = \frac{\Gamma(n+k+1)}{\Gamma(n+\alpha+1)} \sum_{m=0}^{k} (-1)^m \binom{k}{m} L_{n+k-m}^{(\alpha-k+m)}(x)\, L_{n+m}^{(\alpha-m)}(u)\ .^*$$

On prouve sans peine que $G_n^{(\alpha,k)}$ vérifie la relation

$$(u-x)\,G_n^{(\alpha,k)} = G_n^{(\alpha,k+1)} + 2kG_{n+1}^{(\alpha-1,k-1)} - G_{n-1}^{(\alpha,k+1)}$$

d'où cette autre relation

$$(16)\qquad (u-x)\sum_{m=n_1}^{m=n_2} G_m^{(\alpha,k)} = G_{n_2}^{(\alpha,k+1)} - G_{n_1}^{(\alpha,k+1)} + 2k\sum_{n_1+1}^{n_2+1} G_m^{(\alpha-1,k-1)}.$$

Pour $n < 0$ on constate, en posant $n = -p$ que $G_{-p}^{(\alpha,k)} \equiv 0$ pour $2p > k$.
On a en particulier

$$G_{-k}^{(\alpha,2k)} = \frac{(-1)^k(2k)!}{k!\,\Gamma(\alpha-k+1)} \text{ et } G_{-k}^{(\alpha,2k+1)} = \frac{(-1)^k(2k+1)!\,(u-x)}{k!\,\Gamma(\alpha-k+1)}.$$

Vu que $n!\,L_n^{(\alpha)}(x)\,L_n^{(\alpha)}(u) = G_n^{(\alpha,0)}\Gamma(n+\alpha+1)$ on déduit facilement les expressions:

$$\sigma_n^{(\alpha,0)} = (u-x)^{-1}G_n^{(\alpha,1)}$$

$$\sigma_n^{(\alpha,1)} = (u-x)^{-2}G_n^{(\alpha,2)} + 2(u-x)^{-2}\sigma_{n+1}^{(\alpha-1,0)}$$

$$\sigma_n^{(\alpha,2)} = (u-x)^{-3}G_n^{(\alpha,3)} + 6(u-x)^{-2}\sigma_{n+1}^{(\alpha-1,1)}$$

où l'on a posé $\sigma_{-p}^{(\alpha,k)} \equiv 0$ pour $p = [p] \geqq 1$.

En général, on a pour $k \geqq 0$ $n = [n]$

$$(17)\qquad \sigma_n^{(\alpha,k)} = (u-x)^{-(k+1)}G_n^{(\alpha,k+1)} + \sum_{m=1}^{\left[\frac{k+1}{2}\right]} c_m^{(k)}(u-x)^{-2m}\sigma_{n+m}^{(\alpha-m,k-m)}.$$

* On pose $L_{-q}^{(\alpha)}(x) = 0$ pour $q = [q] \geqq 1$.

Les coefficients numériques $c_m^{(k)}$ ne dépendent ni de n ni de α. Pour vérifier (17) supposons la établie déjà pour $k = 0, 1, 2, \cdots l-1, l$ et prouvons qu'elle est exacte alors aussi pour $k = l+1$.

À cet effet posons dans (17) $k = l$ donnons à n les valeurs $0, 1, 2, \cdots n-1, n$ et sommons. On trouve pour

$$\sigma_n^{(\alpha, l+1)} = \sigma_0^{(\alpha, l)} + \sigma_1^{(\alpha, l)} + \cdots + \sigma_n^{(\alpha, l)}$$

une expression qui se ramène à l'aide de (16) et de (17), d'où l'on tire $G_n^{(\alpha-1, l)}$ en fonction des sigma-sommes, à

$$\sigma_n^{(\alpha, l+1)} = (u-x)^{-(l+2)} G_n^{(\alpha, l+2)} + \sum_{m=1}^{\left[\frac{l+2}{2}\right]} c_m^{(l+1)} (u-x)^{-2m} \sigma_{n+m}^{(\alpha-m, l-m+1)} - \rho_n .$$

Dans cette expression de $\sigma_n^{(\alpha, l+1)}$ on a posé

$$c_1^{(l+1)} = c_1^{(l)} + 2(l+1)$$

$$c_{1+[l/2]}^{(l+1)} = -[1 + (-1)^l](l+1) c_{[l/2]}^{(l-1)}$$

ainsi que pour $2 \leqq m \leqq \left[\frac{l+1}{2}\right]$:

$$c_m^{(l+1)} = c_m^{(l)} - 2(l+1) c_{m-1}^{(l-1)}$$

et enfin

$$\rho_n = (u-x)^{-(l+2)} \cdot G_{-1}^{(\alpha, l+2)} + \sum_{m=1}^{\left[\frac{l+2}{2}\right]} c_m^{(l+1)} (u-x)^{-2m} \sigma_{m-1}^{(\alpha-m, l-m+1)} .$$

Par conséquent pour achever la vérification de la relation (17) avec $k = l+1$, il suffit de prouver qu'elle est exacte pour $k = l+1$ et $n = -1$, car cela entraine $\rho_n = 0$. Supposons d'abord que l est pair, $l = 2\lambda$, et formons $\sigma_{-p}^{(\alpha, 2\lambda)}$ pour $1 \leqq p \leqq \lambda + 1$. Toutes ces expressions sont nulles et, en les sommant, on trouve d'après (17) :

$$0 = \sum_{p=1}^{\lambda+1} (u-x)^{-(2\lambda+1)} G_{-p}^{(\alpha, 2\lambda+1)} + \sum_{m=1}^{\lambda} c_m^{(2\lambda)} (u-x)^{-2m} \sigma_{m-1}^{(\alpha-m, 2\lambda-m+1)}$$

car

$$\sum_{p=1}^{\lambda+1} \sigma_{m-p}^{(\alpha-m, 2\lambda-m)} = \sigma_{m-1}^{(\alpha-m, 2\lambda-m+1)} .$$

D'autre part, on a

$$G_{-1}^{(\alpha,k+2)} = G_{-(q+1)}^{(\alpha,k+2)} - 2(k+1)\sum_{p=0}^{q-1} G_{-p}^{(\alpha-1,k)} + (u-x)\sum_{p=1}^{q} G_{-p}^{(\alpha,k+1)}.$$

En y posant $k = 2\lambda$ et $q = \lambda + 1$, on en déduit vu que $G_{-(\lambda+2)}^{(\alpha,2\lambda+2)} = 0$:

$$(u-x)\sum_{p=1}^{\lambda+1} G_{-p}^{(\alpha,2\lambda+1)} = G_{-1}^{(\alpha,2\lambda+2)} + 2(2\lambda+1)\sum_{p=0}^{\lambda} G_{-p}^{(\alpha-1,2\lambda)}$$

donc

$$\sum_{m=1}^{\lambda} c_m^{(2\lambda)}(u-x)^{-2m}\sigma_{m-1}^{(\alpha-m,2\lambda-m+1)}$$

$$= -(u-x)^{-(2\lambda+2)}\left\{G_{-1}^{(\alpha,2\lambda+2)} + 2(2\lambda+1)\sum_{p=0}^{\lambda} G_{-p}^{(\alpha-1,2\lambda)}\right\}.$$

Formant maintenant $\sigma_{-p}^{(\alpha-1,2\lambda-1)}$ pour $p = 0, 1, 2, \cdots \lambda$ et sommant, on trouve de même

$$\sigma_0^{(\alpha-1)} = (u-x)^{-2\lambda}\sum_{m=0}^{\lambda} G_{-m}^{(\alpha-1,2\lambda)} + \sum_{m=2}^{\lambda+1} c_{m-1}^{(2\lambda-1)}(u-x)^{2-2m}\sigma_{m-1}^{(\alpha-m,2\lambda-m+1)}.$$

A l'aide de ces relations on vérifie facilement que $\rho_n = 0$. Le cas d'un l impair, $l = 2\lambda + 1$, se laisse traiter de la même manière et nous pouvons considérer la relation (17) comme vérifiée.

Supposons maintenant que l'inégalité

$$\sigma_n^{(\alpha,k)}(x,u) = O\left\{\frac{e^{\frac{x+u}{2}}(ux)^{-\frac{\alpha}{2}-\frac{1}{4}}\sqrt{n^k}}{|\sqrt{u}-\sqrt{x}|^{k+1}}\right\} \tag{18}$$

soit démontrée pour $k = 0, 1, 2, \cdots l-1$ et vérifions qu'alors elle est exacte aussi pour $k = l$. L'expression de $G_n^{(\alpha,k)}$ donne à l'aide de l'inégalité (2) §2

$$G_n^{(\alpha,k)} = O\left[n^{\frac{k-1}{2}} e^{\frac{x+u}{2}}(ux)^{-\frac{\alpha}{2}-\frac{1}{4}}\sum_{m=0}^{k}\binom{k}{m}x^{\frac{k-m}{2}}u^{\frac{m}{2}}\right]$$

$$= O\left[e^{\frac{x+u}{2}}(\sqrt{u}+\sqrt{x})^k(ux)^{-\frac{\alpha}{2}-\frac{1}{4}}\sqrt{n^{k-1}}\right]$$

d'où vérification facile de (18) pour $k = 0, 1$ pourvu que l'on ait $|u-x|\sqrt{n} \geqq 1$.

La relation (17) donne, supposant (18) valable pour $0 \leqq k \leqq l-1$

$$|\sigma_n^{(\alpha,l)}(x,u)| \leqq \frac{|G_n^{(\alpha,l+1)}|}{|u-x|^{l+1}} + \sum_{m=1}^{\left[\frac{l+1}{2}\right]} |c_m^{(l)}| e^{\frac{x+u}{2}} \frac{(u-x)^{-2m} O\left[(n+m)^{\frac{l-m}{2}}\right]}{(ux)^{\frac{\alpha-m}{2}+\frac{1}{4}} |\sqrt{u}-\sqrt{x}|^{l-m+1}}$$

$$= O\left\{\frac{e^{\frac{x+u}{2}}\sqrt{n^l}}{(ux)^{\frac{\alpha}{2}+\frac{1}{4}}|\sqrt{u}-\sqrt{x}|^{l+1}} \left[1 + \sum_{1}^{\left[\frac{l+1}{2}\right]} |c_m^{(l)}| \frac{\left(\frac{\sqrt{ux}}{\sqrt{u}+\sqrt{x}}\right)^m}{(|u-x|\sqrt{n})^m}\right]\right\}$$

$$= O\left\{\frac{e^{\frac{x+u}{2}}(ux)^{-\frac{\alpha}{2}-\frac{1}{4}}\sqrt{n^l}}{|\sqrt{u}-\sqrt{x}|^{l+1}}\right\}.$$

Ainsi, la preuve de (18) est faite et pour δ entier l'inégalité

$$L_n^{(\delta)}(x,u) = O\left\{\frac{e^{\frac{x+u}{2}}(ux)^{-\frac{\alpha}{2}-\frac{1}{4}}}{|\sqrt{u}-\sqrt{x}|^{\delta+1}\sqrt{n^\delta}}\right\}$$

est prouvée quelque soit u dans l'intervalle $(0, \infty)$ sous l'unique condition $|\sqrt{u}-\sqrt{x}|\sqrt{n} \geqq 1$.

Il est d'ailleurs facile de lever cette restriction. On a évidemment à l'aide de la formule

$$L_n^{(0)}(x,u) = \frac{\Gamma(n+2)}{\Gamma(n+\alpha+1)} \frac{L_{n+1}^{(\alpha-1)}(u)\,L_n^{(\alpha)}(x) - L_{n+1}^{(\alpha-1)}(x)\,L_n^{(\alpha)}(u)}{x-u}$$

l'évaluation surivante basée sur l'inégalité (2) §2.

$$|L_n^{(0)}(x,u)| \leqq O\left\{n^{l-\alpha} \frac{e^{\frac{x+u}{2}} n^{\frac{\alpha-1}{2}-\frac{1}{4}+\frac{\alpha}{2}-\frac{1}{4}}}{(ux)^{\frac{\alpha}{2}+\frac{1}{4}}|u-x|}(\sqrt{u}+\sqrt{x})\right\}$$

c'est à dire

$$L_n^{(0)}(x,u) = O\left(\frac{e^{\frac{x+u}{2}}(ux)^{-\frac{\alpha}{2}-\frac{1}{4}}}{|\sqrt{u}-\sqrt{x}|}\right).$$

A fortiori on aura pour tout $\delta \geqq 0$ entier ou non, et dans l'hypothèse $|\sqrt{u} - \sqrt{x}|\sqrt{n} < 1$

$$|L_n^{(\delta)}(x,u)| \leqq \max_{0 \leq m \leq n} |L_n^{(0)}(x,u)| = O\left(\frac{e^{\frac{x+u}{2}}(ux)^{-\frac{\alpha}{2}-\frac{1}{4}}}{|\sqrt{u}-\sqrt{x}|}\right)$$

$$= O\left\{\frac{e^{\frac{x+u}{2}}(ux)^{-\frac{\alpha}{2}-\frac{1}{4}}}{|\sqrt{u}-\sqrt{x}|\,(|\sqrt{u}-\sqrt{x}|\sqrt{n})^{\delta}}\right\}.$$

Ceci posé, supposons que l'on ait $\eta n \leqq u \leqq e^2 n$. L'inégalité à démontrer s'écrit, vu que x est fini et positif, $x > 0$:

$$\sigma_n^{(\delta)}(x,u) = O\left[u^{-\frac{\alpha}{2}-\frac{3}{4}} e^{\frac{u}{2}} \left(\frac{n}{u}\right)^{\frac{\delta}{2}}\right] = O\left[e^{\frac{u}{2}} u^{-\frac{\alpha}{2}-\frac{3}{4}}\right]$$

et l'on sait qu'elle est vérifiée pour tout δ entier. Soit donc $\delta \neq [\delta]$ et posons $\delta = k - f$, où $k = [\delta + 1]$ et $0 < f < 1$. On peut écrire:

$$|\sigma_n^{(\delta)}(x,u)| \leqq \sum_0^n |A_{n-m}^{(-f-1)} \sigma_m^k|$$

$$= O\left\{u^{-\frac{\alpha}{2}-\frac{3}{4}} e^{\frac{u}{2}} \sum_0^{\infty} (m+1)^{-(f+1)}\right\} = O\left[u^{-\frac{\alpha}{2}-\frac{3}{4}} e^{\frac{u}{2}}\right].$$

Ainsi la preuve de l'inégalité (AL) dans l'intervalle considéré est faite. Quant à celle de (BL), observons qu'elle est absolument analogue à la preuve de la même inégalité pour la série (H) exposée au §3 c'est pourquoi nous l'omettons.

Soit maintenant $0 \leqq u \leqq u_0$, où u_0 est un nombre aussi grand qu'on veut, mais fixe. La formule approchée de M. Maitland Wright s'applique et l'on ne doit tenir compte que d'un seul point singulier $z = +1$. En effet, pour la fonction

$$\sum_0^{\infty} z^n \sigma_n^{(\delta)}(x,u) = \frac{e^{-\frac{(x+u)z}{1-z}} \sqrt{(uxz)^{-\alpha}}}{(1-z)^{\delta+2}} I_{\alpha}\left(\frac{2\sqrt{uxz}}{1-z}\right)$$

l'origine $z = 0$ est un point d'holomorphie et pour $z \to 1$ on a

$$z^{-\frac{\alpha}{2}} I_{\alpha}\left(\frac{2\sqrt{uxz}}{1-z}\right) = \frac{e^{\frac{2\sqrt{ux}}{1-z}} \sqrt{1-z}}{2\sqrt{\pi}\sqrt{ux}}\,\Theta(z).$$

La fonction $\Theta(z)$ n'est pas holomorphe au point $z = 1$, mais pour $z \to 1$ $\Theta(z) \to \Theta(1) = e^{-\sqrt{ux}}$ et l'on a

$$\Theta(z) = e^{-\sqrt{ux}} + O(|1 - z|).$$

En modifiant légèrement l'analyse de M. Maitland Wright, on constate que son résultat reste applicable dans notre cas et de

$$\sum_0^\infty \sigma_n^{(\delta)}(x, u)\, z^n = \frac{e^{x+u}}{2\sqrt{\pi}\,(ux)^{\frac{\alpha}{2}+\frac{1}{4}}} \frac{e^{-\frac{(\sqrt{u}-\sqrt{x})^2}{1-z}}\Theta(z)}{(1-z)^{\delta+\frac{3}{2}}}$$

on déduit la formule approchée suivante:

$$\sigma_n^{(\delta)}(x, u) = \frac{e^{x+u-\sqrt{ux}}}{4\pi(ux)^{\frac{\alpha}{2}-\frac{1}{4}}}\, e^{\frac{\sigma}{2}e^{i\theta}}(\sigma e^{i\theta})^{\frac{1}{4}-l} n^{l-\frac{3}{4}}$$

$$\times\left\{e^{2e^{i\frac{\theta}{2}}\sqrt{\sigma n}} + e^{-2e^{i\frac{\theta}{2}}\sqrt{\sigma n}+i\pi\left(2l-\frac{1}{2}\right)} + O\left(n^{-\frac{1}{2}}\right)\right\}$$

où $\theta = \pi$, $\sigma = (\sqrt{u} - \sqrt{x})^2$ et $l = \dfrac{\delta}{2} + \dfrac{3}{4}$, c'est à dire

$$\frac{\sigma_n^{(\delta)}}{A_n^{(\delta)}} = L_n^{(\delta)}(x, u)$$

$$(19)\qquad = \frac{\Gamma(\delta+1)\, e^{\frac{u+x}{2}}\left\{\cos\left[2\,|\sqrt{u}-\sqrt{x}|\sqrt{n} - \dfrac{\delta+1}{2}\pi\right] + O\left(n^{-\frac{1}{2}}\right)\right\}}{2\pi\,(ux)^{\frac{\alpha}{2}+\frac{1}{4}}\,|\sqrt{u}-\sqrt{x}|^{\delta+1}\sqrt{n^\delta}}.$$

Cette formule approchée prouve que, pour u et x finis tous les deux, on a l'inégalité (AL).

Pour former la formule approchée pour $\dfrac{\partial \sigma_n^{(\delta)}(u, x)}{\partial x}$ observons que l'on a:

$$(20)\qquad \frac{I'_\alpha(\tau)}{I_\alpha(\tau)} = 1 - \frac{1}{2\tau} + \frac{\alpha^2 - \frac{1}{4}}{2\tau^2}\left(1 + \frac{1}{\tau}\right) + O(|\tau|^{-4})$$

d'où, en posant $\tau = \dfrac{2\sqrt{uxz}}{1-z}$:

$$\frac{(1-z)\dfrac{\partial}{\partial x}\Psi_\alpha^{(\delta)}(z)}{\Psi_\alpha^{(\delta)}(z)} = \frac{I'_\alpha(\tau)}{I_\alpha(\tau)}\sqrt{\frac{uz}{x}} - z - \frac{\alpha(1-z)}{2x}.$$

Le second membre tend pour $z \to 1$ vers $\sqrt{\frac{u}{x}} - 1$ et il est égal à:

$$\frac{\sqrt{u} - \sqrt{x}}{\sqrt{x}} + O(|1 - z|)$$

donc pour $z \to 1$ on a

$$\sum_0^\infty z^n \frac{\partial}{\partial x} \sigma_n^{(\delta)}(x, u) = \frac{\sqrt{u} - \sqrt{x}}{\sqrt{x}} \frac{e^{x+u-\sqrt{ux}}}{2\sqrt{\pi}(ux)^{\frac{\alpha}{2}+\frac{1}{4}}} \frac{e^{-\frac{(\sqrt{u}-\sqrt{x})^2}{1-z}} \vartheta(z)}{(1-z)^{\delta+\frac{5}{2}}}$$

où $$\vartheta(z) = 1 + O(|1 - z|).$$

La formule approchée s'écrit donc ainsi:

$$\frac{\partial \sigma_n^{(\delta)}(u, x)}{\partial x} = \frac{\sqrt{u} - \sqrt{x}}{\sqrt{x}} \cdot \frac{e^{\frac{u+x}{2}} \left\{\cos\left[2|\sqrt{u} - \sqrt{x}|\sqrt{n} - \frac{\delta\pi}{2} - \pi\right] + O\left(n^{-\frac{1}{2}}\right)\right\} n^{\frac{\delta+1}{2}}}{2\pi(ux)^{\frac{\alpha}{2}+\frac{1}{4}} |\sqrt{u} - \sqrt{x}|^{\delta+2}}$$

ce qui donne

$$\frac{\partial L_n^{(\delta)}(x, u)}{\partial x} \tag{21}$$
$$= -\frac{\Gamma(\delta + 1) e^{\frac{u+x}{2}} \sqrt{n}\left\{\cos\left[2|\sqrt{u} - \sqrt{x}|\sqrt{n} - \frac{\delta\pi}{2}\right] + O\left(n^{-\frac{1}{2}}\right)\right\}}{2\pi u^{\frac{\alpha}{2}+\frac{1}{4}} x^{\frac{\alpha}{2}+\frac{3}{4}} |\sqrt{u} - \sqrt{x}|^{\delta+1} \sqrt{n^\delta}}$$

et enfin l'inégalité cherchée (BL).

§7

Pour achever la démonstration des inégalités fondamentales relatives à $L_n^{(\delta)}(x, u)$ et $\frac{\partial}{\partial x} L_n^{(\delta)}(x, u)$ il nous reste à évaluer ces moyennes dans l'intervalle $u_0 < u < \eta n$, où u_0 et η sont deux quantités positives choisies d'avance donc fixes, u_0 étant aussi grand et η aussi petit qu'on veut. A cet effet nous allons appliquer la méthode du col à l'intégrale définie

$$\frac{1}{2\pi i}\int_{(C)} F(z) \cdot dz = \frac{1}{2\pi i}\int_{(C)} \frac{e^{-\frac{u+x}{1-z}}}{z^n(1-z)^\beta} I_\alpha\left(\frac{2\sqrt{uxz}}{1-z}\right) dz = J_n(\alpha, \beta; x, u),$$

où le nombre n (entier ou non) tend vers l'infini. En ce qui concerne α et β soient $\alpha > -1$ et $\beta \geqq 2$ vu que le système des polynômes de Laguerre, n'est orthogonal que pour $\alpha > -1$ et que β, égal dans notre cas à $\delta + 2$ ou à $\delta + 3$, est sûrement non inférieur à deux pour $\delta \geqq 0$.

En formant l'équation $F'(z) = 0$ et en utilisant la relation (20) on constate que cette équation s'écrit

$$\frac{F'(z)}{F(z)} = (4n-1)z^3 - 2(4n - 2x - 2u - 2\beta)z^2 + (4n - 4\beta + 1)z$$

$$- 4(1+z)z\sqrt{uxz} - \frac{4\alpha^2 - 1}{16ux}(1+z)(1-z)^2(1 - z + 2\sqrt{uxz})$$

$$+ O\left(\frac{|1-z|^4}{u\sqrt{u}}\right) = 0\,.$$

Les hypothèses

$$\frac{1}{\epsilon} < (\sqrt{u} - \sqrt{x})^2 < (\sqrt{u} + \sqrt{x})^2 < \epsilon n$$

où ϵ est une quantité aussi petite qu'on veut choisie d'avance, déterminent le choix des u_0 et η et avec ce choix, on trouve les deux cols z_0 et $\bar{z}_0$ en fonction du rapport $h = \dfrac{i(\sqrt{u} - \sqrt{x})}{\sqrt{4n - 2\beta}}$, qui est très petit, $|h| < \sqrt{\epsilon}$. Notons encore pour abréger $\sqrt{u} - \sqrt{x} = d > 0$ et nous avons, en omettant les calculs très longs mais élémentaires et faciles à reconstituer

$$z_0 = \rho_0 e^{i\theta_0} = x_0 + iy_0 = \varphi(h) = 1 + 2h + 2h^2 + h^3$$

$$+ (2\beta - 1)\,d^{-2}h^2 - d^{-1}h^3\sqrt{x} - \tfrac{3}{2}h^5 - 2\,d^{-1}h^4\sqrt{x}$$

$$+ (4\beta - 2 - x)\,d^{-2}h^3 - \tfrac{5}{4}h^6 - 9\,d^{-1}h^5\sqrt{x}$$

$$+ (5\beta - 2x - \tfrac{5}{2})\,d^{-2}h^4 - (\alpha^2 - \tfrac{1}{4})\,d^{-3}h^3x^{-1} + O\left(\epsilon^{\frac{7}{2}}\right).$$

Le second col, $\bar{z}_0$, conjugué au premier est donné par $\bar{z}_0 = x_0 - iy_0 = \varphi(-h)$. On trouve en partant de ce résultat les expressions suivantes qui nous seront utiles dans la suite:

$$\theta_0 = \frac{d}{\sqrt{n}}\left\{1 + \frac{d^2 + 3d\sqrt{x}}{24\,n} + O(\epsilon^2)\right\} < \sqrt{\epsilon},$$

$$\rho_0 = |\bar{z}_0| = 1 - \frac{2\beta - 1}{n} + O(\epsilon^3) < 1,$$

$$x_0 = 1 - \frac{2d^2 + 2\beta - 1}{4n} - \frac{d^3\sqrt{x}}{8n^2} + O(\epsilon^3) < 1,$$

$$y_0 = \frac{d}{\sqrt{n}}\left\{1 - \frac{d^2 - d\sqrt{x} - x + 4\beta - 1}{8n} - \frac{3d^4}{64n^2} + O\left(\epsilon^{\frac{5}{2}}\right)\right\} < \sqrt{\epsilon}.$$

Calculons encore la distance $r_0 = |1 - z_0|$ du col z_0 au point $z = 1$:

$$r_0 = |1 - z_0| = \frac{d}{\sqrt{n}}\left\{1 + \frac{d\sqrt{x} + x - 4\beta + 1}{8n} - \frac{5d^4}{128n^2} + O\left(\epsilon^{\frac{5}{2}}\right)\right\}.$$

Nous définissons maintenant le chemin d'intégration (C), qui entoure l'origine, ainsi : soit θ l'angle que fait avec OX la ligne de la plus grande pente passant par le col z_0. Les cols sont à l'intérieure du cercle $|z| = 1$. Soient C et E les points de sa circonférence, où les lignes $z = z_0 + te^{i\vartheta}$ et $z = \bar{z}_0 + te^{-i\vartheta}$ coupent le cercle $|z| = 1$. Soient de même B et F les points de l'intersection de ces lignes avec le cercle concentrique de rayon $\rho_1 < \rho_0$ le nombre ρ_1 devant être déterminé dans la suite. Désignons enfin par ω et par θ_1 les angles que font avec OX les rayons vecteurs des points B et C. Le chemin (C) se compose:

1°) de l'arc $\overset{\frown}{FAB}$ du cercle de rayon ρ_1,

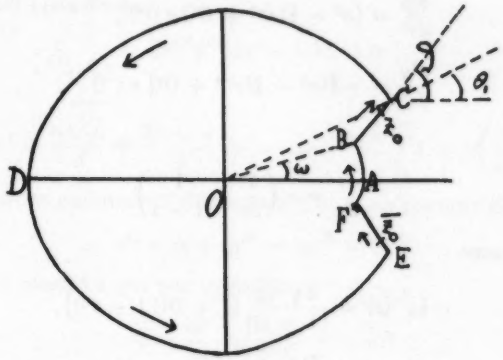
2°) du segment $\overline{BC}$ passant par le col z_0 et tangent au point z_0 à la ligne de la plus grande pente,

3°) de l'arc $\overset{\frown}{CDE}$ du cercle de rayon 1 et,

4°) du segment $\overline{EF}$ passant par le col $\bar{z}_0$ et symétrique au segment $\overline{CB}$. Sur le segment $\overline{BC}$ on a

$$z = z_0 + te^{i\vartheta}.$$

Soient $t_1 < 0$ et $t_2 > 0$ les valeurs de paramètre t correspondant aux points B et C.



On voit immédiatement que l'on a

$$\pi J_n(\alpha, \beta; x, u) = \pi J_n = i_1 + i_2 + i_3$$

$$= \rho_1 \int_0^{\omega} R[e^{i\varphi} F(\rho_1 e^{i\varphi})]\, d\varphi + \int_{t_1}^{t_2} R[e^{i\left(\vartheta - \frac{\pi}{2}\right)} F(z_0 + te^{i\vartheta})]\, dt$$

$$+ \int_{\theta_1}^{\pi} R[e^{i\varphi} F(e^{i\varphi})]\, d\varphi .$$

Pour déterminer ϑ posons $F(z) = F(z_0)\, e^{\psi(z)}$ et $z = z_0 + te^{i\vartheta}$. L'angle ϑ est déterminé par la condition que la quantité $e^{2i\vartheta}\psi''(z_0)$ est réelle et négative.

On a $F'(z_0) = 0$ donc $\psi(z_0) = \psi'(z_0) = 0$ et $F(z_0)\psi''(z_0) = F''(z_0)$. Pour déterminer le rapport

$$N = -\tfrac{1}{2}\frac{F''(z_0)}{F(z_0)} = -\frac{n}{2z_0^2} - \frac{\beta}{2(1-z_0)^2} + \frac{u+x}{(1-z_0)^3} - \tfrac{1}{2}\, y''(z_0)$$

où

$$y(z) = \frac{\tau I_\alpha'(\tau)}{I_\alpha(\tau)} \quad \text{avec} \quad \tau = \frac{2\sqrt{uxz}}{1-z},$$

observons que l'on a

$$\frac{dy}{d\tau} = 1 - \frac{4\alpha^2 - 1}{8\tau^2}\left(1 - \frac{1}{\tau}\right) + O(|\tau|^{-4}),$$

$$\frac{d^2y}{d\tau^2} = (\alpha^2 - \tfrac{1}{4})\tau^{-2} + O(|\tau|^{-4}),$$

$$\frac{d^3y}{d\tau^3} = -3(\alpha^2 - \tfrac{1}{4})\tau^{-4} + O(|\tau|^{-6}).$$

Or,

$$\frac{d\tau}{dz} = \tau\left(\frac{1}{2z} + \frac{1}{1-z}\right)$$

ce qui entraine

$$\tfrac{1}{2}y''(z) = \frac{2\sqrt{ux}}{(1-z)^3}[1 + O(|1-z|)],$$

$$\tfrac{1}{6}y'''(z) = \frac{2\sqrt{ux}}{(1-z)^4}[1 + O(|1-z|)].$$

Etant donné que $|1 - z_0| = r_0 = dn^{-\frac{1}{2}}[1 + O(n^{-1})]$ on trouve

$$N = -\tfrac{1}{2}\psi''(z_0) = \frac{d^2}{(1-z_0)^3} - \frac{n}{2z_0^2} + O\left(\frac{d}{r_0^2}\right).$$

Si l'on pose $1 - z_0 = r_0 e^{-i\chi_0}$ on trouve

$$\chi_0 = \frac{\pi}{2} - \frac{d}{\sqrt{n}} + O(\epsilon).$$

Par conséquent:

$$N = \frac{d^2}{r_0^3}e^{3i\chi_0} - \frac{\bar{n}}{2}e^{-2i\vartheta_0} + O\left(\frac{n}{d}\right) = -\frac{7n}{2}[1 + O(\sqrt{\epsilon})] - \frac{in\sqrt{n}}{d}[1 + O(\epsilon)].$$

On en déduit, en exigeant que le produit $Ne^{2i\vartheta}$ soit réel et positif

$$\frac{2\sqrt{n}}{7d}\cos 2\vartheta + \sin 2\vartheta = O\left(\frac{1}{d}\right)$$

donc

$$\operatorname{tang} 2\vartheta = -\frac{2\sqrt{n}}{7d}\left[1 + O\left(n^{-\frac{1}{2}}\right)\right]$$

et

$$\vartheta = \frac{\pi}{4} + \frac{7d}{4\sqrt{n}}\left[1 + O\left(\frac{1}{\sqrt{n}}\right)\right].$$

Ce résultat nous donne

$$\tfrac{1}{2}e^{2i\vartheta}\psi''(z_0) = -\lambda^2 < 0$$

avec

$$\lambda^2 = \frac{n\sqrt{n}}{d} + \frac{49\sqrt{n}}{4} + \cdots = \frac{n\sqrt{n}}{d}\left[1 + O\left(\frac{d}{n}\right)\right].$$

Nous pouvons maintenant déterminer t_2 et θ_1 en partant de la relation

$$e^{i\theta_1} = z_0 + t_2 e^{i\vartheta} = \rho_0 e^{i\theta_0} + t_2 e^{i\vartheta}.$$

On obtient ainsi, vu que par hypothèse $d^{-1} < \sqrt{\epsilon}$:

$$t_2 = \frac{2\beta - 1}{2n\sqrt{2}}\left[1 + O(\sqrt{\epsilon})\right]$$

et

$$\theta_1 = \frac{d}{\sqrt{n}}\left\{1 + \frac{d^2 + 3d\sqrt{x}}{24n} + \frac{2\beta - 1}{4d\sqrt{n}} + O(\epsilon^2)\right\}.$$

Quant à t_1 et ω on les déterminera, en évaluant l'intégrale i_2 en valeur absolue:

$$|\, i_2 \,| < \int_{t_1}^{t_2} |\, F(z_0 + te^{i\vartheta}) \,|\, dt$$

où

$$|\, F(z_0 + te^{i\vartheta}) \,| \leqq |\, F(z_0) \,|\, e^{-\lambda^2 t^2 + \frac{t^3}{6} R\left[e^{3i\vartheta}\psi'''(z_0 + hte^{i\vartheta})\right]}$$

la valeur $t' = ht$ du paramétre t étant comprise entre 0 et t, c'est à dire $0 < h < 1$. A fortiori on a, en posant $F(z_0) = F_0$

$$|\, i_2 \,| < |\, F_0 \,| \int_{t_1}^{t_2} e^{-\lambda^2 t^2 + \frac{t^3}{6}|\psi'''(\zeta)|}\, dt,$$

où ζ est un point du segment $(z_0, z_0 + te^{i\vartheta})$. Or,

$$\psi'''(z) = -2nz^{-3} + 2\beta(1 - z)^{-3} - 6(u + x)(1 - z)^{-4} + y'''(z)$$

$$= -2nz^{-3} - 6d^2(1 - z)^{-4}\left[1 + O\left(\frac{|\, 1 - z \,|}{d}\right)\right],$$

car

$$y'''(z) = \frac{12\sqrt{ux}}{(1 - z)^4}\left[1 + O(|1 - z\,|)\right] = \frac{12\sqrt{ux}}{(1 - z)^4} + O\left(\frac{d}{|\, 1 - z \,|^3}\right).$$

Par conséquent sur le segment $\overline{BC}$ on a sûrement

$$|\psi'''(z)| \leqq \frac{6d^2}{|1-z|^4}\left\{1 + O\left(\frac{|1-z|}{d}\right)\right\} + O(n).$$

Pour $z = z_0 + te^{i\vartheta}$ le minimum de $r = |1 - z|$ correspond à

$$t = t_m = (1 - x_0)\cos\vartheta - y_0\sin\vartheta \sim -d(2n)^{-\frac{1}{2}}$$

et ce minimum r_m est égal à

$$r_m = (1 - x_0)\sin\vartheta + y_0\cos\vartheta \sim \frac{d}{\sqrt{2n}}$$

d'où la conclusion:

$$\frac{t^3}{6}|\psi'''(\zeta)| \leqq \frac{4n^2t^3}{d^2}[1 + O(\epsilon)]$$

Soit maintenant

$$\lambda^2t^2 \geqq \frac{8n^2t^3}{d^2}[1 + O(\epsilon)]$$

ce qui entraîne, grâce au fait que $\lambda^2 \sim \frac{n\sqrt{n}}{d}$:

$$|t| \leqq \frac{d}{8\sqrt{n}}. \tag{22}$$

Dans ces conditions on voit qu'il faut poser $t_1 = -\frac{d}{8\sqrt{n}}$, la valeur de $t_2 = O(n^{-\frac{1}{4}})$ vérifiant sûrement la condition (22).

Par conséquent on a dans l'intervalle (t_1, t_2)

$$|F(z_0 + t\,e^{i\vartheta})| < |F_0|\,e^{-\frac{1}{2}\lambda^2t^2}$$

et l'on peut conclure:

$$|i_2| \leqq |F_0|\int_{t_1}^{t_2} e^{-\frac{\lambda^2}{2}t^2}\,dt < \frac{|F_0|\sqrt{2}}{\lambda}\int_{-\infty}^{\infty} e^{-w^2}\,dw < \frac{\pi\sqrt{2d}}{\sqrt[4]{n^3}}\,|F_0|.$$

Or,

$$|F_0| = |z_0^{-n}\,(1 - z_0)^{-\beta}\,e^{-\frac{u+z}{1-z}}\,I_\alpha(\tau_0)|$$

$$= \frac{r_0^{-\beta}\,\rho_0^{-n}}{\sqrt{2\pi\,|\tau_0|}}\,e^{R\left[\tau_0 - \frac{u+z}{1-z_0}\right]}\left\{1 + O\left(\frac{1}{|\tau_0|}\right)\right\}$$

car $\tau_0 = \frac{2\sqrt{uxz_0}}{1 - z_0} = \frac{2\sqrt{xu}}{1 - z_0} - \frac{2\sqrt{xu}}{1 + \sqrt{z_0}}$ est très grand en valeur absolue.

insi

$$|F_0| = \frac{\rho_0^{-n-\frac{1}{4}}\, r_0^{\frac{1}{2}-\beta}}{2\sqrt{\pi}\,\sqrt[4]{ux}}\, e^{-R\left(\frac{d^2}{1-z_0}+\frac{2\sqrt{xu}}{1+\sqrt{z_0}}\right)}\left\{1+O\left(\frac{1}{|\tau_0|}\right)\right\}.$$

Or, pour $\beta > \frac{1}{2}$ on peut écrire

$$R\left(\frac{2\sqrt{xu}}{1+\sqrt{z_0}}\right) = \frac{2\sqrt{ux}\left(1+\sqrt{\rho_0}\,\cos\frac{\theta_0}{2}\right)}{1+2\sqrt{\rho_0}\,\cos\frac{\theta_0}{2}+\rho_0}$$

$$= \sqrt{ux}\left(1+\frac{2\beta-1}{16n}+\cdots\right) > \sqrt{ux}\,.$$

De même

$$R\left(\frac{d^2}{1-z_0}\right) = \frac{d^2\, r_0\cos\chi_0}{r_0^2} = \frac{d^2\,(1-x_0)}{r_0^2} = \frac{d^2}{2}+\frac{2\beta-1}{4}$$

$$+\frac{d^2}{2}\left\{\frac{5\,d^4}{128n^2}-\frac{3(2\beta-1)}{4n}+O\left(\epsilon^{\frac{5}{2}}\right)\right\}.$$

Si l'on a $d \geqq 2\,\sqrt[4]{1.2\;n(2\beta-1)}$ on trouve

$$e^{-R\left[\frac{d^2}{1-z_0}\right]} \leqq e^{-\frac{d^2}{2}-\frac{2\beta-1}{4}}\,[1+O(\epsilon)]\,.$$

Mais même si l'on a

$$d < 2\,\sqrt[4]{1.2\;n\,(2\beta-1)}$$

on en tire

$$\frac{d^2}{2}\left\{\frac{3(2\beta-1)}{4n}-\frac{5d^4}{128n^2}+O\left(\epsilon^{\frac{5}{2}}\right)\right\} = O\left(\frac{d^6}{n^2}\right) = O\left(\frac{1}{\sqrt{n}}\right) = O(\epsilon)$$

et par conséquent la même conclusion subsiste et l'on a toujours:

$$e^{-R\left\{\frac{d^2}{1-z_0}+\frac{2\sqrt{u\,x}}{1+\sqrt{z_0}}\right\}} \leqq e^{-\frac{x+u}{2}-\frac{2\beta-1}{4}}\,[1+O(\epsilon)]\,.$$

En outre, vu que $r_0\sqrt{n} > d$ et $\beta \geqq 2$, on constate que

$$|1-z_0|^{\frac{1}{2}-\beta} = r_0^{\frac{1}{2}-\beta} < \left(\frac{d}{\sqrt{n}}\right)^{\frac{1}{2}-\beta}$$

et

$$\rho_0^{-\left(n+\frac{1}{4}\right)} \leqq \left(1 - \frac{2\beta - 1}{4n}\right)^{-\left(n+\frac{1}{4}\right)} = e^{\frac{2\beta-1}{4}} \left[1 + O\left(\frac{1}{n}\right)\right]$$

ce qui donne le résultat suivant

$$| F_0 | < \frac{e^{-\frac{x+u}{2}}}{2\sqrt{\pi}\sqrt[4]{ux}} \left(\frac{\sqrt{n}}{d}\right)^{\beta-\frac{1}{2}} [1 + O(\epsilon)]$$

et enfin

$$| i_2 | < \frac{| F_0 | \sqrt{2d}}{\sqrt[4]{n^3}} < \frac{e^{-\frac{x+u}{2}} n^{\frac{\beta}{2}-1} [1 + O(\epsilon)]}{\sqrt{2\pi}\sqrt[4]{ux} | \sqrt{u} - \sqrt{x} |^{\beta-1}}.$$

Passons au calcul de la borne supérieure pour l'intégrale

$$i_1 = \rho_1 \int_0^{\omega} R\{e^{i\varphi} F(\rho_1 e^{i\varphi})\} d\varphi .$$

On a

$$| i_1 | < \int_0^{\omega} | F(\rho_1 e^{i\varphi}) | d\varphi ,$$

car

$$\rho_1 e^{i\omega} = z_0 + t_1 e^{i\vartheta} = \rho_0 e^{i\theta_0} - \frac{d e^{i\vartheta}}{8\sqrt{n}}$$

donne:

$$\rho_1 = 1 - \frac{d}{8\sqrt{2n}} + O(\epsilon)$$

$$\omega = \frac{8\sqrt{2} - 1}{8\sqrt{2n}} d \left[1 - \frac{3(129 + 16\sqrt{2})}{2032\sqrt{2}} \frac{d}{\sqrt{2}} + O(\epsilon)\right].$$

Pour $z = \rho_1 e^{i\vartheta}$, $0 \leqq \varphi \leqq \omega < \frac{d}{\sqrt{n}}$ on trouve

$$| F(\rho_1 e^{i\vartheta}) | \leqq \frac{e^{n - \frac{u+x}{2}} \rho_1^{-n} (1 - \rho_1)^{\frac{1}{2}-\beta}}{2\sqrt{\pi}\sqrt[4]{ux\rho_1}} \left\{1 + O\left(\frac{1}{|\tau|}\right)\right\}$$

où

$$\frac{1}{|\tau|} = O\left(\frac{|1 - z|}{d}\right) = O\left(\frac{\sqrt{1 - 2\rho_1 \cos\omega + \rho_1^2}}{d}\right) = O\left(\frac{1}{\sqrt{n}}\right) = O(\epsilon)$$

et

$$\mu = R\left(\tau - \frac{x+u}{1-z}\right) + \frac{x+u}{2} < \frac{d^2\,(1-\rho_1^2)}{2\,[1-2\rho_1\cos\varphi+\rho_1^2]}.$$

Or,

$$1-2\rho_1\cos\varphi+\rho_1^2 = (1-\rho_1)^2 + 4\rho_1\sin^2\frac{\varphi}{2}$$

$$< \frac{d^2}{128n}[1+O(\epsilon)] + \omega^2 < \frac{27d^2}{32n}[1+O(\epsilon)]$$

et par conséquent

$$\frac{d^2(1-\rho_1^2)}{2\,(1-2\rho_1\cos\varphi+\rho_1^2)} > \frac{32n}{27d^2}\,\frac{d^2}{2}\,\frac{d}{4\sqrt{2n}}\,[1+O(\sqrt{\epsilon})]$$

d'où

$$\mu < -\frac{d\sqrt{n}}{7\sqrt{2}}\,\{1+O(\sqrt{\epsilon})\}.$$

En outre

$$\rho_1^{-n} = \left[\left(1-\frac{d}{8\sqrt{2n}}\right)^{-\sqrt{n}}\right]^{\sqrt{n}} \leqq e^{\frac{d\sqrt{n}}{8\sqrt{2}}[1+O(\sqrt{\epsilon})]}$$

ce qui nous donne

$$e^{\mu}\,\rho_1^{-n} \leqq e^{-\frac{d\sqrt{n}}{56\sqrt{2}}[1+o(\sqrt{\epsilon})]} < e^{-c_1 d\sqrt{n}}$$

où c_1 est une constante positive, $c_1 > 0$, et enfin

$$|F(\rho_1 e^{i\varphi})| < O\left\{e^{-\frac{x+u}{2}}\;n^{\frac{\beta}{2}-\frac{1}{4}}\cdot d^{-\beta}\;e^{-c_1 d\sqrt{n}}\right\}$$

quelque soit φ dans l'intervalle $(0, \omega)$. Ainsi, on a trouvé que pour $n \to \infty$

$$|i_1| < |i_2|\,e^{-c_2\sqrt{n}} \qquad (c_2 > 0)$$

où la constante c_2 est positive, $c_2 > 0$.

Considérons maintenant

$$i_3 = \int_{\theta_1}^{\pi} R\{e^{i\varphi}\,F(e^{i\varphi})\}\,d\varphi,$$

où

$$e^{i\varphi} F(e^{i\varphi}) = e^{-i(n-1)\varphi - \frac{x+u}{1-e^{i\varphi}}} (1 - e^{i\varphi})^{-\beta} I_\alpha\left(\frac{i\sqrt{ux}}{\sin\frac{\varphi}{2}}\right)$$

$$= e^{-i\left(n+\frac{\beta}{2}-1\right)\varphi + \frac{(x+u)e^{-i\frac{\varphi}{2}}}{2i\sin\frac{\varphi}{2}}} \left(-2i\sin\frac{\varphi}{2}\right)^{-\beta} J_\alpha\left(\frac{\sqrt{ux}}{\sin\frac{\varphi}{2}}\right) i^\alpha.$$

Par conséquent, on trouve

$$R(e^{i\varphi} F(e^{i\varphi})) = \left(2\sin\frac{\varphi}{2}\right)^{-\beta} e^{-\frac{x+u}{2}} \cos[\Omega(\varphi)]\, J_\alpha\left(\frac{\sqrt{ux}}{\sin\frac{\varphi}{2}}\right)$$

où

$$\Omega(\varphi) = \frac{x+u}{2}\cot\frac{\varphi}{2} + \left(n + \frac{\beta}{2} - 1\right)\varphi - (\alpha+\beta)\frac{\pi}{2}.$$

Mais

$$J_\alpha\left(\frac{\sqrt{ux}}{\sin\frac{\varphi}{2}}\right) = \frac{\sqrt{2\sin\frac{\varphi}{2}}}{\sqrt{\pi}\sqrt[4]{ux}} \cos\left[\frac{\sqrt{ux}}{\sin\frac{\varphi}{2}} - \left(\alpha + \frac{1}{2}\right)\frac{\pi}{2}\right] + O\left\{\left(\frac{\varphi}{\sqrt{ux}}\right)^{\frac{3}{2}}\right\}$$

et pour $\beta > 2$ cela entraîne, en notant, pour abréger, l'argument du cosinus par $\Psi(\varphi)$:

$$2^{\frac{x+u}{2}} i_3 = \frac{1}{2\sqrt{\pi}\sqrt[4]{ux}} \int_{\theta_1}^{\pi} \frac{2\cos\Omega(\varphi)\cos\Psi(\varphi)\, d\varphi}{\left(2\sin\frac{\varphi}{2}\right)^{\beta-\frac{1}{2}}} + j_3$$

où

$$\Psi(\varphi) = \frac{\sqrt{ux}}{\sin\frac{\varphi}{2}} - \left(\alpha + \frac{1}{2}\right)\frac{\pi}{2} \qquad (\beta > 2)$$

et

$$j_3 = O\left\{(ux)^{-\frac{3}{4}} \int_{\theta_1}^{\pi} \varphi^{\frac{3}{2}-\beta}\, d\varphi\right\},$$

l'angle θ_1 étant égal à

$$\theta_1 = \frac{d}{\sqrt{n}}\left[1 + \frac{d^2 + 3d\sqrt{x}}{24n} + \frac{2\beta - 1}{4d\sqrt{n}} + O(\epsilon^2)\right].$$

Pour $\beta > \frac{5}{2}$ on trouve, vu que $\sqrt{u} \sim d$ et que $\sqrt{d} \leqq \sqrt[4]{n}\epsilon$:

$$j_3 = O\left[\theta_1^{\frac{5}{2}-\beta}\ (ux)^{-\frac{3}{4}}\right] = n^{\frac{\beta}{2}-1}\ (ux)^{-\frac{1}{4}}\ d^{1-\beta}\ O(\sqrt[4]{\epsilon})$$

tandis que, si $2 < \beta \leqq \frac{5}{2}$, on peut écrire

$$j_3 = O\left[(xu)^{-\frac{3}{4}}\ |\log \theta_1|\right] = n^{\frac{\beta}{2}-1}\ (xu)^{-\frac{1}{4}}\ d^{1-\beta}\ O\left(\epsilon^{\frac{\beta}{2}-1}\ |\log \epsilon|\right).$$

Ainsi pour tout $\beta > 2$ on trouve

$$j_3 = o\left[n^{\frac{\beta}{2}-1}\ (xu)^{-\frac{1}{4}}\ d^{1-\beta}\right] \qquad (\beta > 2).$$

Si $\beta = 2$, on doit employer la formule approchée

$$J_\alpha\left(\frac{\sqrt{ux}}{\sin\frac{\varphi}{2}}\right) = \frac{\sqrt{2\sin\frac{\varphi}{2}}}{\sqrt{\pi}\ \sqrt[4]{ux}}\left\{\cos[\Psi(\varphi)] - \frac{4\alpha^2 - 1}{8\sqrt{ux}}\sin\frac{\varphi}{2}\ \sin[\Psi(\varphi)]\right\}$$
$$+ O\left\{\left(\frac{\varphi}{\sqrt{ux}}\right)^{\frac{5}{2}}\right\}.$$

Par conséquent pour $\beta = 2$ nous avons

$$j_3 = -\frac{4\alpha^2 - 1}{32\sqrt{\pi}\ (ux)^{3/4}}\int_{\theta_1}^{\pi}\frac{2\cos\Omega\ \sin\Psi}{\sqrt{2\sin\frac{\varphi}{2}}}\,d\varphi + O\left[(ux)^{-\frac{5}{4}}\right].$$

Laissons pour le moment ce cas $\beta = 2$ de coté et occupons nous du cas général $\beta > 2$. Dans l'expression de la partie principale de i_3 intervient la somme des deux intégrales

$$\int_{\theta_1}^{\pi}\frac{\cos(\Omega + \Psi)}{\left(2\sin\frac{\varphi}{2}\right)^{\beta-\frac{1}{2}}}\,d\varphi + \int_{\theta_1}^{\pi}\frac{\cos(\Omega - \Psi)}{\left(2\sin\frac{\varphi}{2}\right)^{\beta-\frac{1}{2}}}\,d\varphi = i_4 + i_5$$

auxquelles nous allons appliquer le lemme du §1.

Les deux arguments des cosinus sont:

$$U(\varphi) = \Omega + \Psi = \frac{x+u}{2}\cot\frac{\varphi}{2} + \frac{\sqrt{ux}}{\sin\dfrac{\varphi}{2}} + \left(n + \frac{\beta}{2} - 1\right)\varphi - \left(2\alpha + \beta + \frac{1}{2}\right)\frac{\pi}{2},$$

$$V(\varphi) = \Omega - \Psi = \frac{x+u}{2}\cot\frac{\varphi}{2} - \frac{\sqrt{ux}}{\sin\dfrac{\varphi}{2}} + \left(n + \frac{\beta}{2} - 1\right)\varphi - \frac{\pi}{2}\left(\beta - \frac{1}{2}\right).$$

La dérivée de $U(\varphi)$

$$U'(\varphi) = n + \frac{\beta}{2} - 1 - \frac{x + 2\cos\dfrac{\varphi}{2}\sqrt{ux} + u}{4\sin^2\dfrac{\varphi}{2}}$$

s'annule pour $\varphi = \varphi^*$ où, en posant $\sqrt{u} + \sqrt{x} = s$:

$$\cos\frac{\varphi^*}{2} = 1 - \frac{s^2}{8n} - \frac{d^2 s^2}{128n^2} - \frac{(2-\beta)s^2}{16n^2} + \frac{d^4 s^2}{1024\, n^3} + O(\epsilon^4)$$

et $U'(\varphi) \lesseqgtr 0$ suivant que $\varphi \lesseqgtr \varphi^*$. En calculant $\sin\dfrac{\varphi^*}{4}$ et ensuite φ^*, on trouve la formule approchée

$$\varphi^* = \frac{s}{\sqrt{n}}\left[1 + O(\sqrt{\epsilon})\right].$$

On a en outre $U''(\varphi) > 0$, car

$$4\sin^3\frac{\varphi}{2}\, U'''(\varphi) = (x+u)\cos\frac{\varphi}{2} + \left(1 + \cos^2\frac{\varphi}{2}\right)\sqrt{ux}$$

et

$$-8\sin^4\frac{\varphi}{2}\, U''''(\varphi) = (x+u)(2+\cos\varphi) + \left(5 + \cos^2\frac{\varphi}{2}\right)\cos\frac{\varphi}{2}\sqrt{ux}$$

donc $U''''(\varphi) < 0$. Considérons maintenant l'intégrale

$$i_6 = \int_{\varphi^*}^{\pi} \frac{\cos\left[U(\varphi)\right] d\varphi}{\left(2\sin\dfrac{\varphi}{2}\right)^{\beta - \frac{1}{2}}}$$

où $U' > 0$, $U'' > 0$, $U''' < 0$, la fonction

$$F(\varphi) = \left(2 \sin \frac{\varphi}{2}\right)^{\frac{1}{2}-\beta}$$

étant décroissante, et appliquons le lemme du §1.

On a ici, vu que $F'(\varphi) < 0$, $M = F(\varphi^*)$ et $c = \varphi^*$, car $U'' > 0$. Eu égard à $U'(\varphi^*) = 0$ et au fait que $U''' < 0$, on trouve:

$$|i_6| < \frac{F(\varphi^*)\sqrt{2\pi}}{\sqrt{U''(\varphi^*)}} .$$

Or, l'inégalité facile à démontrer

$$4 \sin^3 \frac{\varphi}{2} U''(\varphi) \geqq s^2 \cos \frac{\varphi}{2}$$

prouve que l'on a pour $\beta > 2$

$$|i_6| < \frac{\left(2 \sin \frac{\varphi^*}{2}\right)^{2-\beta} \sqrt{\pi}}{\sqrt[s]{\cos \frac{\varphi^*}{2}}} < \left(\frac{\pi}{2}\right)^{\beta-\frac{3}{2}} \frac{n^{\frac{\beta}{2}-1}}{d^{\beta-1}} \sqrt{2}\,[1 + O(\epsilon)]$$

car

$$\varphi^* \sqrt{n} = s\,[1 + O(\epsilon)] > d\,[1 + O(\epsilon)] .$$

En calculant la différence $\cos \frac{\theta_1}{2} - \cos \frac{\varphi^*}{2}$ on trouve:

$$\cos \frac{\theta_1}{2} - \cos \frac{\varphi^*}{2} = \frac{\sqrt{ux}}{2n} \{1 + O(\epsilon)\} \tag{23}$$

par conséquent on a toujours $\varphi^* > \theta_1$ et nous devons encore considérer l'intégrale étendue à l'intervalle (θ_1, φ^*). Or, (23) nous donne, vu que

$$\varphi^* + \theta_1 = \frac{s + d}{\sqrt{n}} [1 + O(\epsilon)] = 2 \sqrt{\frac{u}{n}} [1 + O(\epsilon)]$$

la relation:

$$\varphi^* = \theta_1 + 4 \sqrt{\frac{x}{n}} [1 + O(\epsilon)] .$$

Ceci posé, considérons l'intgérale

$$i_7 = \int_{\theta_1}^{\varphi^*} \frac{\cos [- U(\varphi)]}{\left(2 \sin \frac{\varphi}{2}\right)^{\beta-\frac{1}{2}}} d\varphi$$

où $f(\varphi) = -U(\varphi)$ est croissante dans (θ_1, φ^*), $f'(\varphi) > 0$, $f''(\varphi) < 0$ et $f'''(\varphi) > 0$. Ceci permet d'appliquer le lemme et, vu que $f'(c) = f'(\varphi^*) = 0$ et

$$M = \left(2 \sin \frac{\theta_1}{2}\right)^{\frac{1}{2}-\beta} = \left(\frac{d}{\sqrt{n}}\right)^{\frac{1}{2}-\beta} [1 + O(\epsilon)]$$

on obtient, tenant compte du fait que $d^2\epsilon > 1$, le résultat suivant:

$$| i_7 | < M \frac{\sqrt{2\pi}}{\sqrt{| f''(\varphi^*) |}} = \sqrt{\pi}\, n^{\frac{\beta}{2}-1}\, d^{1-\beta} [1 + O(\sqrt{\epsilon})] .$$

On parvient ainsi à l'inégalité cherchée pour i_4:

$$| i_4 | < | i_6 | + | i_7 | \leqq \left\{\sqrt{\pi} + \sqrt{2}\left(\frac{\pi}{2}\right)^{\beta-\frac{3}{2}}\right\} n^{\frac{\beta}{2}-1}\, d^{1-\beta} [1 + O(\sqrt{\epsilon})] .$$

Passons à l'intégrale

$$i_5 = \int_{\theta_1}^{\pi} \frac{\cos [V(\varphi)]\, d\varphi}{\left(2 \sin \frac{\varphi}{2}\right)^{\beta-\frac{1}{2}}}$$

pour laquelle

$$4 \sin^2 \frac{\varphi}{2} V'(\varphi) = 4 (n + \beta - 2) \sin^2 \frac{\varphi}{2} + 2 \cos \frac{\varphi}{2} \sqrt{ux} - u - x .$$

Cette fois-ci on a $V'(\varphi_*) = 0$, où

$$\cos \frac{\varphi_*}{2} = 1 - \frac{d^2}{n} - \frac{d^2 s^2}{128 n^2} + O(\epsilon^3) .$$

Par conséquent, la comparaison avec $\cos \frac{\theta_1}{2}$ donne

$$\cos \frac{\varphi_*}{2} - \cos \frac{\theta_1}{2} = \frac{(2\beta - 1)d}{16 n \sqrt{n}} [1 + O(\sqrt{\epsilon})] > 0$$

et ainsi $\theta_1 > \varphi_*$, ce qui prouve que $V'(\varphi)$ ne s'annule pas dans l'intervalle (θ_1, π) et y est constamment positive: $V'(\varphi) > 0$. Ensuite

$$4 \sin^3 \frac{\varphi}{2} V''(\varphi) = d^2 - 2(u + x) \sin^2 \frac{\varphi}{4} + \sin^2 \frac{\varphi}{2} \sqrt{ux} .$$

La dérivée seconde $V''(\varphi)$ s'annule pour $\varphi = \bar{\varphi}$, où

$$\cos \frac{\bar{\varphi}}{2} = \sqrt{\frac{x}{u}}$$

ce qui correspond à $\bar{\varphi} = \pi - \chi$ avec

$$\sin\frac{\chi}{2} = \cos\frac{\bar{\varphi}}{2} = \sqrt{\frac{x}{u}}$$

donc $\chi = \dfrac{2\sqrt{x}}{d}[1 + O(\sqrt{\epsilon})]$. Il est évident que

$$i_8 = \int_{\pi-\chi}^{\pi} \frac{\cos V(\varphi)\, d\varphi}{\left(2\sin\frac{\varphi}{2}\right)^{\beta-\frac{1}{2}}} = \frac{(-1)^n \sqrt{x}\cos(\sqrt{ux})}{2^{\beta-\frac{3}{2}}\sqrt{u}}[1 + O(\sqrt{\epsilon})]$$

et par conséquent pour $\beta > 2$

$$|i_8| < \left(\frac{\sqrt{n}}{d}\right)^{\beta-2} \epsilon^{\frac{\beta}{2}-1} \frac{\sqrt{x}}{\sqrt{u}\, 2^{\beta-\frac{3}{2}}}[1 + O(\sqrt{\epsilon})] = n^{\frac{\beta}{2}-1} d^{1-\beta} O\left(\epsilon^{\frac{\beta}{2}-1}\right).$$

Quant à l'intervalle $(\theta_1, \pi - \chi)$ on a dans cet intervalle $V'(\varphi) > 0$, $V''(\varphi) > 0$ et $V'''(\varphi) < 0$. Par conséquent, on peut appliquer le lemme à l'intégrale

$$i_9 = \int_{\theta_1}^{\pi-\chi} \frac{\cos[V(\varphi)]\, d\varphi}{\left(2\sin\frac{\varphi}{2}\right)^{\beta-\frac{1}{2}}}$$

avec $c = \theta_1$, ce qui nous donne le résultat:

$$|i_9| < M\left\{\left(\left[\frac{V_1'}{V_1''}\right]^2 + \frac{2\pi}{V_1''}\right)^{\frac{1}{2}} - \frac{V_1'}{V_1''}\right\} = Mh$$

avec

$$M = \left(2\sin\frac{\theta_1}{2}\right)^{\frac{1}{2}-\beta} = \left(\frac{d}{\sqrt{n}}\right)^{\frac{1}{2}-\beta}[1 + O(\epsilon)].$$

En calculant les valeurs V_1' et V_1'' des dérivées $V'(\varphi)$ et $V''(\varphi)$ pour $\varphi = \theta_1$, on trouve

$$V_1' = \left(\beta - \frac{1}{2}\right)\frac{\sqrt{n}}{d}[1 + O(\sqrt{\epsilon})]$$

tandis que

$$V_1'' = \frac{2n\sqrt{n}}{d}[1 + O(\epsilon)].$$

Par conséquent le rapport $V_2'^3\, V_1''^{-1}$ est très petit:

$$\frac{V_1'^3}{V_1''} = \frac{(2\beta-1)^2}{32d\sqrt{n}}\,[1+O(\sqrt{\epsilon})]$$

et cela nous permet d'écrire

$$h = \left\{\left(1+\frac{V_1'^3}{2\pi V_1''}\right)^{\frac{1}{2}} - \frac{V_1'}{\sqrt{2\pi V_1''}}\right\}\sqrt{\frac{2\pi}{V_1''}} = \frac{\sqrt{\pi d}}{\sqrt[4]{n^3}}\,[1+O(\sqrt[4]{\epsilon})]$$

et par conséquent

$$|\, i_9\,| < n^{\frac{\beta}{2}-1}\, d^{1-\beta}\,\sqrt{\pi}\,[1+O(\sqrt[4]{\epsilon})]$$

ce qui achève la preuve de l'inégalité

$$|\, i_8\,| < n^{\frac{\beta}{2}-1}\, d^{1-\beta}\,\sqrt{\pi}\,[1+o(1)] \qquad (\beta > 2)\,.$$

Il nous reste à vérifier seulement que la dérivée tierce $V'''(\varphi)$ est négative. On a:

$$8\sin^4\frac{\varphi}{2}\,V'''(\varphi) = \left(5+\cos^2\frac{\varphi}{2}\right)\cos\frac{\varphi}{2}\,\sqrt{ux} - (2+\cos\varphi)\,(u+x)$$

et, en étudiant le second membre pour $0 \leqq \varphi \leqq \pi$, on trouve qu'il reste constammant négatif dans l'intervalle $(0, \pi)$ et même dans celui $(0, 2\pi)$.

Le cas $\beta > 2$ est ainsi étudiée et l'on trouve que dans ce cas on a:

$$e^{\frac{x+u}{2}}\,|\, i_3\,| \leqq \frac{|\, i_4\,| + |\, i_5\,|}{2\sqrt{\pi}\,\sqrt[4]{ux}} + |\, j_2\,|$$

$$\leqq n^{\frac{\beta}{2}-1}\, d^{1-\beta}\,(ux)^{-\frac{1}{4}}\left[1+\frac{1}{2}\left(\frac{\pi}{2}\right)^{\beta-2}+o(1)\right]$$

c'est à dire pour $\beta > 2$ on a demontré l'inégalité

$$i_3 = O\left\{\frac{e^{-\frac{x+u}{2}}\; n^{\frac{\beta}{2}-1}}{\sqrt[4]{ux}\; d^{\beta-1}}\right\} \qquad (\beta \geqq 2)\,. \tag{24}$$

Soit maintenant $\beta = 2$. L'hypothèse $\beta > 2$ n'ayant servi que lors du calcul de la borne supérieure de l'intégrale i_8 et de celle de j_3 nous ne devons considérer que ces deux quantités. Quant à i_8 on a:

$$|\, i_8\,| < \sqrt{\frac{x}{2u}}\,[1+O(\sqrt{\epsilon})] \qquad (\beta = 2)\,.$$

Considérons maintenant

$$j_3 = \frac{4\alpha^2 - 1}{32\sqrt{\pi}\sqrt[4]{(ux)^3}}(j_4 + j_5) + O\left(u^{-\frac{5}{4}}\right)$$

où l'on a posé

$$j_4 = -\int_{\theta_1}^{\pi} \frac{\sin U(\varphi)\, d\varphi}{\sqrt{2\sin\frac{\varphi}{2}}}, \qquad j_5 = \int_{\theta_1}^{\pi} \frac{\sin V(\varphi)\, d\varphi}{\sqrt{2\sin\frac{\varphi}{2}}}.$$

En appliquant aux deux intégrales j_4 et j_5 le lemme, on obtient:

$$|j_4| < 2\sqrt{\frac{\pi}{n}}[1 + o(1)], \qquad j_5 < \sqrt{\frac{\pi}{n}} + \sqrt{\frac{2x}{u}}.$$

Par conséquent, on trouve pour $|j_3|$ la borne supérieure

$$j_3 = O\left(\frac{1}{d\sqrt{n}\sqrt[4]{u}} + \frac{1}{u^{\frac{5}{4}}}\right) = \frac{O(\sqrt{\epsilon})}{d\sqrt[4]{u}}.$$

Cela prouve que j_3 rentre dans le reste de l'inégalité

$$e^{\frac{x+u}{2}}|i_3| \doteq \frac{3 + x\sqrt{2}}{2d\sqrt[4]{ux}}[1 + o(1)]$$

qui justifie la validité de (24) aussi pour $\beta = 2$.

Nous pouvons maintenant conclure. Dans l'intervalle $u_0 \leqq u \leqq \eta n$ on obtient l'inégalité fondamentale pour $J_n(\alpha, \beta; u, x)$, où $\alpha > -1$, $\beta \geqq 2$:

$$\pi|J_n| < |i_1| + |i_2| + |i_3|$$

$$< \frac{e^{-\frac{x+u}{2}} n^{\frac{\beta}{2}-1}}{d^{\beta-1}\sqrt[4]{ux}}\left\{\frac{1}{\sqrt{2\pi}} + 1 + \frac{1}{2}\left(\frac{\pi}{2}\right)^{\beta-2} + \frac{x}{\sqrt{2}}\right\}[1 + o(1)]$$

d'où pour x compris dans un intervalle fini, $0 < x \leqq a$:

$$J_n(\alpha, \beta; u, x) = O\left\{e^{-\frac{x+u}{2}} n^{\frac{\beta}{2}-1} d^{1-\beta} (ux)^{-\frac{1}{4}}\right\}.$$

On a, pour $n = m + 1 + \alpha/2$

$$\sum_{m=0}^{\infty} J_{m+1+\alpha/2}\, z^m = \frac{e^{-\frac{u+x}{1-z}}}{z^{\alpha/2}(1-z)^{\beta}} I_{\alpha}\left(\frac{2\sqrt{uxz}}{1-z}\right).$$

Or, pour $\sigma_n^{(\delta)}(x, u) = A_n^{(\delta)} L_n^{(\delta)}(x, u)$ on a la fonction génératrice

$$\sum_{n=0}^{\infty} \sigma_n^{(\delta)}(u, x)\, z^n = \frac{e^{-\frac{(x+u)z}{1-z}}}{\sqrt{(uxz)^\alpha}\ (1-z)^{\delta+2}} I_\alpha\left(\frac{2\sqrt{uxz}}{1-z}\right).$$

Par conséquent, on trouve:

$$\sigma_n^{(\delta)}(u, x) = \frac{e^{x+u}}{\sqrt{(ux)^\alpha}} J_{n+1+\alpha/2}(\alpha, \delta+2; u, x)$$

d'où l'inégalité pour $L_n^{(\delta)}(u, x)$, valable si $\delta \geqq 0$:

$$\text{(AL)} \qquad L_n^{(\delta)}(u, x) = O\left\{\frac{e^{\frac{x+u}{2}}\ n^{-\frac{\delta}{2}}}{(ux)^{\frac{\alpha}{2}+\frac{1}{4}}\ |\sqrt{u} - \sqrt{x}\,|^{\delta+1}}\right\}.$$

Cette inégalité est ainsi établie dans l'intervalle $u_0 \leqq u \leqq \eta\, n$ pour tout x fini, α étant supérieur à, -1 et δ positif ou nul, $\delta \geqq 0$. Ainsi, la preuve de l'inégalité fondamentale (AL), valable dans $0 \leqq u \leqq \infty$ est parachevée.

Passons maintenant à $\frac{\partial}{\partial x} L_n^{(\delta)}(u, x)$. En calculant la fonction génératrice de la suite $\frac{\partial}{\partial x} \sigma_n^{(\delta)}(u, x)$, on trouve

$$\sum_0^\infty z^n \frac{\partial \sigma_n^{(\delta)}(u, x)}{\partial x} = \frac{z\, e^{-\frac{(u+x)z}{1-z}}}{(1-z)^{\delta+3}\sqrt{(ux)^\alpha}} \left\{\frac{I_{\alpha+1}(\tau)}{\sqrt{z^{\alpha+1}}}\sqrt{\frac{u}{x}} - \frac{I_\alpha(\tau)}{\sqrt{z^\alpha}}\right\}$$

où $\tau = \dfrac{2\sqrt{uxz}}{1-z}$, ,

$$\frac{\partial \sigma_n^{(\delta)}(x, u)}{\partial x} = \frac{e^{x+u}}{\sqrt{(ux)^\alpha}} \left\{\sqrt{\frac{u}{x}}\, J_{n+\frac{\alpha+1}{2}}(\alpha+1, \delta+3; u, x) - J_{n+\alpha/2}(\alpha, \delta+3; u, x)\right\}$$

À l'aide de (24) on en déduit

$$\frac{\partial \sigma_n^{(\delta)}(x, u)}{\partial x} = O\left\{\frac{e^{\frac{x+u}{2}}}{(ux)^{\frac{\alpha}{2}+\frac{1}{4}}}\left[\frac{n^{\frac{\delta+1}{2}}}{d^{2+\delta}}\sqrt{\frac{u}{x}} + \frac{n^{\frac{\delta+1}{2}}}{d^{2+\delta}}\right]\right\}$$

c'est à dire après la division par $A_n^{(\delta)} \sim n^\delta\,[\Gamma(\delta+1)]^{-1}$

$$\text{(BL)} \qquad \frac{\partial L_n^{(\delta)}(x,u)}{\partial x} = O\left\{\frac{e^{\frac{x+u}{2}}\; n^{\frac{1-\delta}{2}}}{u^{\frac{\alpha}{2}+\frac{1}{4}}\; x^{\frac{\alpha}{2}+\frac{3}{4}}\;|\sqrt{u}-\sqrt{x}\,|^{\delta+1}}\right\}$$

pour tout $x > 0$ et $u_0 \leqq u \leqq \eta\; n$.

Ainsi, la preuve de l'inégalité fondamentale (BL) dans l'intervalle $0 \leqq u \leqq \infty$ est achevée.

Observons en terminant que les inégalités (AL) et (BL) prouvent que les séries divergentes

$$e^{-\frac{u}{2}} \sum_0^\infty \frac{\Gamma(n+1)}{\Gamma(n+\alpha+1)}\, L_n^{(\alpha)}(x)\, L_n^{(\alpha)}(u) \sim 0 \qquad (C, \delta > 0)$$

et

$$e^{-\frac{u}{2}} \sum_1^\infty \frac{\Gamma(n+1)}{\Gamma(n+\alpha+1)}\, L_{n-1}^{(\alpha+1)}(x)\, L_n^{(\alpha)}(u) \sim 0 \qquad (C, \delta > 1)$$

pour tout u dans $(0, \infty)$ sont uniformément sommables $(C, \delta > 0)$ et $(C, \delta > 1)$ respectivement et.ont leurs sommes nulles pourvu que l'on ait $|u - x| \geqq \epsilon$, x étant fini.

THE LIFT AND THE MOMENT OF AN ARBITRARY AEROFOIL-JOUKOVSKY POTENTIAL

By J. G. Estes[1]

Introduction. The purpose of this paper is to give a method for determining the lift and the moment of an arbitrary aerofoil section, when the flow about the section is two dimensional Joukovsky[2] potential flow. The idea of using the normal variation of Green's function to solve the arbitrary aerofoil problem is due to Prof. Norbert Wiener of the Massachusetts Institute of Technology, under whom the author wrote his doctor's dissertation, for the Department of Mathematics, on the above problem. The present paper introduces some new methods and formulae developed by the author, which simplify the arbitrary aerofoil problem very greatly. This paper will be presented in two parts, the first part being the development of the method and the second part the application of the method developed to a concrete case.

PART I

1. **Outline of the Method.** We can find the lift and the moment of an arbitrary aerofoil, by the use of Blasius' formulae, if we know the potential of the flow about the arbitrary aerofoil. We know the potential of the flow resulting when a circle (cylinder) of radius a with circulation is placed in a flow of uniform velocity.[3] Knowing this, we could get the flow about the arbitrary aerofoil if we knew the complex function which would map the flow about the circle of radius a onto the arbitrary aerofoil, the flow at infinity remaining unchanged.

The following theorem, which is of fundamental importance in this paper, is given in slightly more general form by Kellogg.[4]

If G' is Green's function, for the simply connected domain T, with pole at $\zeta = \infty$, the function $z_1 = ae^{G'+iH'}$, where H' is conjugate to G', maps

[1] Department of Mathematics, North Carolina State College of Agriculture and Engineering.

[2] Glauert, H.; Aerofoil and Airscrew Theory, page 83.

[3] Ibid., page 56.

[4] Kellogg, O. D.; Foundations of Potential Theory, page 366.

T in a one to one, conformal manner on the exterior of the circle $| z_1 | = a$ of the z_1 plane, the point $\zeta = \infty$ going into the point $z_1 = \infty$.

From the preceding theorem, if G' and H' are functions associated with the arbitrary aerofoil, then the function $z_1 = ae^{G'+iH'} = f(\zeta)$, conformally maps the region exterior to the arbitrary aerofoil in the ζ plane onto the region exterior to the circle $| z_1 | = a$ of the z_1 plane, the point $\zeta = \infty$ going into the point $z_1 = \infty$. The inverse of this function, namely, the function which conformally maps the region (and the flow) exterior to the circle $| z_1 | = a$ of the z_1 plane onto the region (and the flow) exterior to the arbitrary aerofoil of the ζ plane, is the conformal mapping function which we are seeking, and the determination of which is our chief concern in this paper.

Because of conformality, the direction at infinity of the flow in the aerofoil plane (ζ plane) will be the same as the direction at infinity of the flow in the z_1 plane. The constant a in the conformal mapping function, it will be seen, is used to make the velocities at infinity in the two planes have the same magnitude.

Our problem thus rests upon finding a certain conformal mapping function, the determination of which depends upon finding G' and H', the Green's function (pole at infinity) and its conjugate function for the region exterior to the arbitrary aerofoil.

In this paper, we shall find G' and H' by variational methods. Let G and H be Green's function (pole at infinity) and its conjugate function for the region exterior to a Joukovsky aerofoil which is near the arbitrary aerofoil. We shall find G' and H' by finding the change in G and the change in H as the near-by Joukovsky aerofoil changes into the arbitrary aerofoil, and adding those changes to G and to H respectively. We thus get G' and H', and then proceed with the solution of the problem, as has here been briefly outlined.

2. Green's Function (G), with Pole at Infinity, and Its Conjugate Function H, for the Region Exterior to a Joukovsky Aerofoil. The problem we are discussing is a problem in conformal mapping, where the flow at infinity must remain unaltered. It is in connection with this fact that lies the reason for the use of Green's functions with their poles at infinity, as is easily seen by referring to the theorem stated in §1.

As stated in §1, we shall need to know Green's function (with pole at infinity) and its conjugate function for the region exterior to a Joukovsky aerofoil; these functions we shall call G and H, respectively.

In this work we shall need to know the conformal transformations which transform the region exterior to a Joukovsky aerofoil into the

region exterior to a circle of radius a, center at the origin. A Joukovsky aerofoil[5] (fig. 1, below) is given by applying the conformal transformation

$$\zeta = z + \frac{c^2}{z} \tag{2.1}$$

to the circle of fig. 2, below. The following figures and conformal transformations show how the region exterior to a Joukovsky aerofoil in the ζ plane is transformed into the region exterior to the circle $|z_1| = a$ of the z_1 plane.

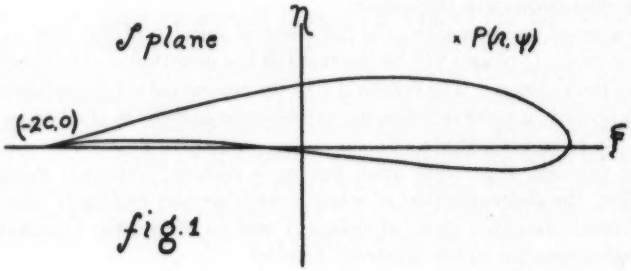
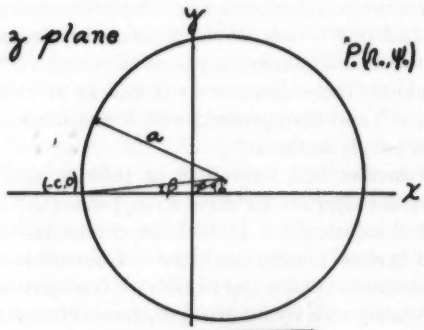


fig. 1



$$z = \frac{\zeta \pm \sqrt{\zeta^2 - 4c^2}}{2}$$

fig. 2

[5] Glauert, H.; Aerofoil and Airscrew Theory, page 71.

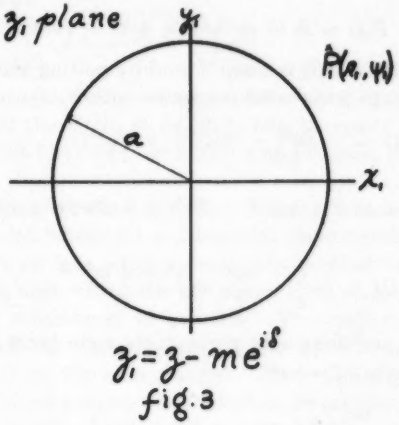


fig. 3

By the theorem stated in §1, and the above transformations;

$$z_1 = ae^{G+iH} = f(\zeta) = \frac{\zeta \pm \sqrt{\zeta^2 - 4c^2} - 2me^{i\delta}}{2} \qquad (2.2)$$

is the mapping function which conformally maps the region exterior to the Joukovsky aerofoil of the ζ plane (fig. 1) onto the region exterior to the circle $| z_1 | = a$ of the z_1 plane, in the desired manner; i.e., the flows at infinity in the ζ and z_1 planes are the same in direction and magnitude. From eq. (2.2), we have

$$G + iH = \ln (\zeta \pm \sqrt{\zeta^2 - 4c^2} - 2me^{i\delta}) - \ln 2a . \qquad (2.3)$$

If we had a series development of the right hand side of the above equation, valid for ζ large, by separating real and imaginary parts, we would have the Fourier series for G and for H. Hence, let $\zeta = \frac{1}{t}$ and expand the right hand side of eq. (2.3) in a power series valid for t small; i.e., valid for those values of t which correspond to values of ζ in the region exterior to the Joukovsky aerofoil of fig. 1. Substituting $\zeta = \frac{1}{t}$ in eq. (2.3), we have

$$G + iH = \ln (1 \pm \sqrt{1 - 4c^2t^2} - 2mte^{i\delta}) - \ln t - \ln 2a . \quad (2.4)$$

Let

$$F(t) = \ln (1 \pm \sqrt{1 - 4c^2t^2} - 2mte^{i\delta}) .$$

The power series for $F(t)$ is easily found by getting the power series for $F'(t)$ and integrating the resulting series term by term.

$$F'(t) = - me^{i\delta} - 2t(c^2 + \tfrac{1}{2} m^2e^{2i\delta}) \ldots\ldots\ldots ,$$

whence

$$F(t) = \ln 2 - mte^{i\delta} - t^2(c^2 + \tfrac{1}{2} m^2e^{2i\delta}) \ldots\ldots\ldots ,$$

and

$$G + iH = \ln \frac{\zeta}{a} - \frac{me^{i\delta}}{\zeta} - \frac{c^2 + \frac{1}{2} m^2e^{2i\delta}}{\zeta^2} \ldots\ldots\ldots$$

Separating real and imaginary parts on the right hand side of the above equation, we have, if $\zeta = \rho e^{i\varphi}$,

$$G(\rho, \varphi) = \ln \frac{\rho}{a} - \frac{m}{\rho} [\cos \delta \cos \varphi + \sin \delta \sin \varphi] - \frac{1}{\rho^2}\left[\left(c^2 + \frac{m^2}{2} \cos 2\delta\right) \cos 2\varphi + \frac{m^2}{2} \sin 2\delta \sin 2\varphi\right] \ldots\ldots\ldots , \tag{2.5}$$

and

$$H(\rho, \varphi) = \varphi - \frac{m}{\rho} [\sin \delta \cos \varphi - \cos \delta \sin \varphi] - \frac{1}{\rho^2}\left[\frac{m^2}{2} \sin 2\delta \cos 2\varphi - \left(c^2 + \frac{m^2}{2} \cos 2\delta\right) \sin 2\varphi\right] \ldots\ldots\ldots \tag{2.6}$$

As we will later use the inverse of Joukovsky's transformation in our calculating, a word about its double sign is in place here. The inverse of Joukovsky's transformation is, from eq. (2.1),

$$z = \frac{\zeta \pm \sqrt{\zeta^2 - 4c^2}}{2} . \tag{2.7}$$

If we place $R + iS = \pm \sqrt{\zeta^2 - 4c^2}$, it can easily be verified that

$$\left.\begin{aligned} R &= \pm \sqrt{\frac{U + E}{2}} \quad \text{and} \quad S = \pm \sqrt{\frac{U - E}{2}} , \\ \text{where} \qquad & \\ U &= + \sqrt{E^2 + 4\xi^2\eta^2} \quad \text{and} \quad E = \xi^2 - \eta^2 - 4 , \end{aligned}\right\} \tag{2.8}$$

and $\zeta = \xi + i\eta$.

Eq. (2.7) becomes

$$z = \frac{(\xi + R) + i(\eta + S)}{2}. \tag{2.9}$$

For the region exterior to the Joukovsky aerofoil of fig. 1 to go into the region exterior to the circle of fig. 2, it can be easily checked that ξ and R must always have like signs, and η and S must always have like signs.

3. **The Determination of G′ and H′.** Suppose we are presented with an arbitrary aerofoil whose lift and moment characteristics are desired. The first thing we do is to select a Joukovsky aerofoil which is near the arbitrary aerofoil, and, with their tail points coinciding, bring the aerofoils as near to coincidence as possible. We shall call the near-by Joukovsky aerofoil the base aerofoil. In fig. 4, let A be the arbitrary aerofoil and let B be the base aerofoil. With the origin as a center, draw the circle C' with a radius equal to about 60 per cent of the common (approximately) chords.

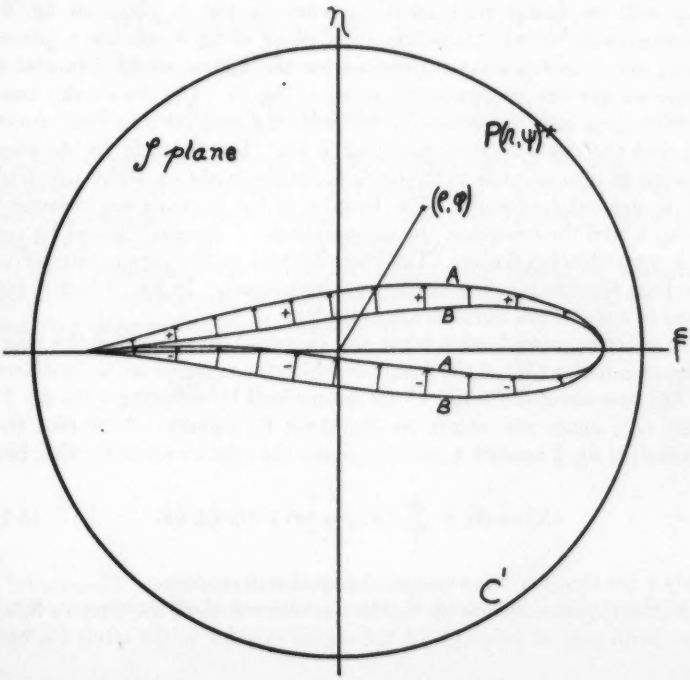


FIG. 4

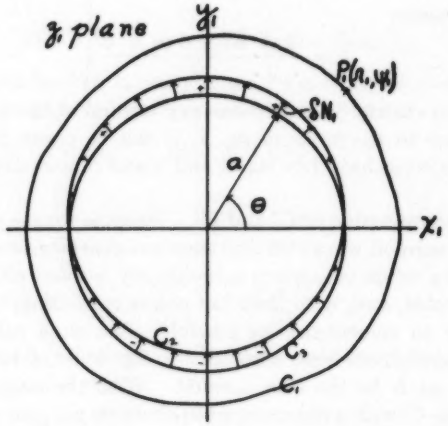


FIG. 5

It will be found convenient to work in the z_1 plane of fig. 3. Consequently, we will transform the ζ plane of fig. 4 into the z_1 plane, using the transformations given under the figures in §2. In the z_1 plane we get the configuration given in fig. 5. The Joukovsky base aerofoil goes into the circle C_3, the arbitrary aerofoil into some curve C_2, and the circle C' into some curve C_1. The normals to the base aerofoil go into normals to the circle C_3, as the transformations involved are in general conformal. The lengths of the normals are measured in fig. 4, and the lengths of the normals in fig. 5 are easily found by use of a magnification factor. The inner normal to the region exterior to the base aerofoil is taken as the positive normal. In figs. 4 and 5, the signs of a few of the normals are indicated.

In practice, considerable labor will be saved in computing the magnification factor (M), if the points at which the normals are to be drawn to the base aerofoil (in fig. 4) are determined by selecting θ (in fig. 5) first; e.g., space the points on the circle C_3 equally. Now plot the normals of fig. 5 against θ, and determine the Fourier series for δN_1; i.e.,

$$\delta N_1 = A_0 + \sum_1^n (A_n \cos n\theta + B_n \sin n\theta) . \tag{3.1}$$

Only a few terms of this series will be needed in practice.

In the z_1 plane, denote by $G_1(P_1)$ the value at P_1 of the Green's function (with pole at infinity) for the region exterior to the circle C_3, and

denote by $H_1(P_1)$ the conjugate function. Using the theorem of §1, we get

$$G_1(P_1) + iH_1(P_1) = \ln \frac{z_1}{a}.$$

From eq. (2.2),

$$G(P) + iH(P) = \ln\left(\frac{\zeta \pm \sqrt{\zeta^2 - 4c^2 - 2me^{i\delta}}}{2a}\right) = \ln \frac{z_1}{a}$$

so, if P and P_1 are transforms of one another, we have the very useful result that;

$$G(P) = G_1(P_1) \text{ and } H(P) = H_1(P_1). \tag{3.2}$$

This result simply says that the value of a transformed function at a transformed point is equal to the value of the original function at the original point. If G_1' and H_1' are, similarly, transformed functions of G' and H', we simply write down, without further discussion, that;

$$G'(P) = G_1'(P_1) \qquad \text{and} \qquad H'(P) = H_1'(P_1). \tag{3.3}$$

Subtracting relations (3.2) from relations (3.3), we get

$$\delta G(P) = \delta G_1(P_1) \text{ and } \delta H(P) = \delta H_1(P_1), \tag{3.4}$$

i.e., the change at P in G and in H as the base aerofoil changes into the arbitrary aerofoil is the same as the change at P_1 in G_1 and H_1 as the circle C_3 changes into the curve C_2. We shall make important use of this result; e.g., we shall get $\delta G(P)$ by calculating $\delta G_1(P_1)$.

Using principles of the Functional Calculus developed by Volterra,[6] the author has derived the following pair of formulae,[7] which we shall find very labor saving in our calculations;

$$\delta G_1(P_1) = -\frac{1}{a}\sum_0^n (A_n X_n + B_n Y_n), \tag{3.5}$$

and

$$\delta H_1(P_1) = +\frac{1}{a}\sum_1^n (A_n Y_n - B_n X_n), \tag{3.6}$$

[6] Volterra, V.; Leçons sur les Fonctions de Lignes.

[7] See a paper by the author in the Journal of Mathematics and Physics, Vol. 13, #2.

where A_n and B_n are from eq. (3.1),

$$X_n = \left(\frac{a}{r_1}\right)^n \cos n\psi_1$$

and

$$Y_n = \left(\frac{a}{r_1}\right)^n \sin n\psi_1 ,$$

where $P_1 = r_1 e^{i\psi_1}$.

Taking several P points on the circle C' of fig. 4 (say at 10° intervals), we calculate the corresponding P_1 points on the curve C_1 of fig. 5, and apply eq. (3.5), getting $\delta G_1(P_1)$ at these P_1 points. But, by the first of relations (3.4), $\delta G(P) = \delta G_1(P_1)$. Thus, we get $\delta G(P)$ by finding $\delta G_1(P_1)$ from eq. (3.5).

Now plot $\delta G(P)$ against ψ (fig. 4), and get the Fourier series for $\delta G(P)$. We find this series to be of the form

$$\delta G(r, \psi) = c_0' + \sum_1^n (c_n' \cos n\psi + d_n' \sin n\psi) . \tag{3.7}$$

We will assume now and justify later, that the constant term c_0' in the above series is independent of r (the radius of C'). The series (3.7) gives the value of δG at points on the circle C' of fig. 4. At any point (ρ, φ), the series (3.7) will take the following form, assuming c_0' is independent of r;

$$\delta G(\rho, \varphi) = c_0 + \sum_1^n (c_n \cos n\varphi + dn \sin n\varphi) \frac{1}{\rho^n} , \tag{3.8}$$

where $c_n' = \frac{c_n}{r^n}$ and $d_n' = \frac{d_n}{r^n}$. In the above series, when $\rho = r$, we get back series (3.7), which is as it should be. The reason for inserting the factor $\frac{1}{\rho^n}$ is easily seen when we recall that the terms of eq. (3.8) are "corrections" to the terms of the series for $G(\rho, \varphi)$, which is written below in slightly different form from that given in eq. (2.5).

$$\begin{aligned} G(\rho, \varphi) = &- \ln a + \ln \rho - \frac{m}{\rho}[\cos \delta \cos \varphi + \sin \delta \sin \varphi] \\ &- \frac{1}{\rho^2}\left[\left(c^2 + \frac{m^2}{2} \cos 2\delta\right) \cos 2\varphi + \frac{m^2}{2} \sin 2\delta \sin 2\varphi\right] + \cdots . \end{aligned} \tag{3.9}$$

Adding series (3.8) and (3.9), we get the series for $G'(\rho, \varphi)$ to be as follows;

$$G'(\rho, \varphi) = \ln \rho + \alpha_0 + \sum_1^n (\alpha_n \cos n\varphi + \beta_n \sin n\varphi) \frac{1}{\rho^n}. \tag{3.10}$$

Next let us justify our assumption regarding c_0' in series (3.7), and at the same time show, by independent considerations, that eq. (3.10) has the correct form. From the theorem of paragraph 1, we have

$$z_1 = ae^{G'+iH'} = f(\zeta) \tag{3.11}$$

from which

$$G' + i\,H' = \ln f(\zeta) - \ln a. \tag{3.12}$$

To investigate the form of the series for the right hand side of eq. (3.12), in the region exterior to the unit circle, let $\zeta = \frac{1}{t}$ and expand the resulting function in the neighborhood of the origin. We have

$$g(t) = \ln f\left(\frac{1}{t}\right) - \ln a = \ln f_1(t) - \ln a. \tag{3.13}$$

When $\zeta = \infty$, $z_1 = \infty$ and hence in eq. (3.11) $G' + iH' = \infty$. In eq. (3.12) $f(\zeta)$ must therefore be infinite when $\zeta = \infty$, and hence $f_1(t)$ in eq. (3.13) is infinite when $t = 0$, and $f_1(t)$ must have the form

$$f_1(t) = \frac{f_2(t)}{t^n}, \tag{3.14}$$

where $f_2(t)$ is analytic and $\neq 0$ at the origin. Using eq. (3.14), eq. (3.13) becomes

$$g(t) = -\ln a - \ln t^n + \ln f_2(t) = -\ln a - n \ln t + b_0 + b_1 t + b_2 t^2 + \cdots$$

where $b_0 \neq 0$.

Replacing t by $1/\zeta$ in the above series, we have

$$G' + iH' = -\ln a + n \ln \zeta + b_0 + \frac{b_1}{\zeta} + \frac{b_2}{\zeta^2} + \cdots.$$

Using eq. (3.11), it may be shown that for conformal mapping at infinity, it is necessary that $n = 1$ in the above series. Taking the real parts of the above series, we find that $G'(\rho, \varphi)$ has the following form;

$$G'(\rho, \varphi) = \ln \rho + E_0 + \sum_1^n (E_n \cos n\varphi + F_n \sin n\varphi) \frac{1}{\rho^n}.$$

The above series is of the same form as series (3.10). This work shows that c_0' of eq. (3.7) is independent of r, as was assumed, because if c_0' were not independent of r, there would be terms in eq. (3.8) involving ρ only, and when eqs. (3.8) and (3.9) were added there would have been terms involving ρ alone in eq. (3.10), other than $\ln \rho$. Incidentally, the above equation also verifies the correctness of inserting the factor $1/\rho^n$ in eq. (3.7) to get eq. (3.8).

In practice it will be convenient to wait until the coefficients in series (3.8) and (3.9) are known, before adding them to get the series (3.10).

From eq. (3.10), we may write down at once the series for $H'(\rho, \varphi)$[8] as;

$$H'(\rho, \varphi) = \beta_0 + \varphi + \sum_1^n (-\alpha_n \sin n\varphi + \beta_n \cos n\varphi) \frac{1}{\rho^n}, \qquad (3.15)$$

with β_0 undetermined. That $\beta_0 = 0$ may be easily seen by observing that eq. (3.15) must approach eq. (2.6) as the arbitrary aerofoil approaches the base aerofoil.

The finding of $G'(\rho, \varphi)$ and $H'(\rho, \varphi)$ is a very important result from the standpoint of this paper, because we are now able, by the theorem of §1, to write down the fundamental mapping function we have been looking for. As has been stated previously, this is the mapping function which maps conformally the region (and the flow) exterior to a circle of radius a on the region (and the flow) exterior to the arbitrary aerofoil. As will be seen in the next paragraph, it will not be necessary to write down this transformation as $\zeta = F(z_1)$, but we will use it in the form $z_1 = ae^{G'+iH'}$. This is our desired conformal mapping function, where $G'(\rho, \varphi)$ and $H'(\rho, \varphi)$ are given by eqs. (3.10) and (3.15). In our calculations, we shall need terms of these series only as far as $n = 2$.

4. **Determination of the Flow about the Arbitrary Aerofoil.** In fig. 6 below is shown the flow resulting, when a circle (cylinder) with circulation is placed in a flow of uniform velocity.

[8] Kellogg, O. D.; Foundations of Potential Theory, page 353.

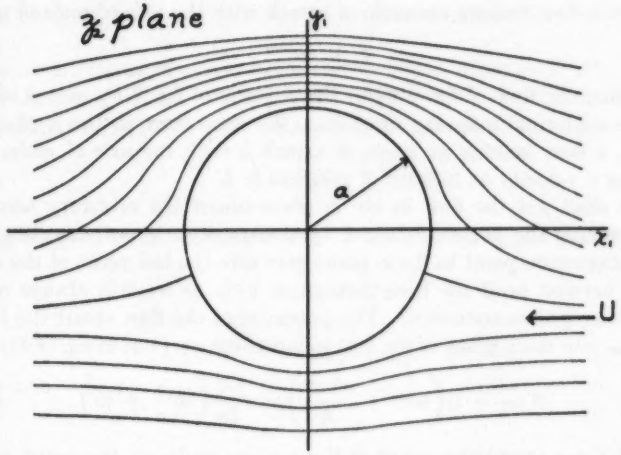


FIG. 6

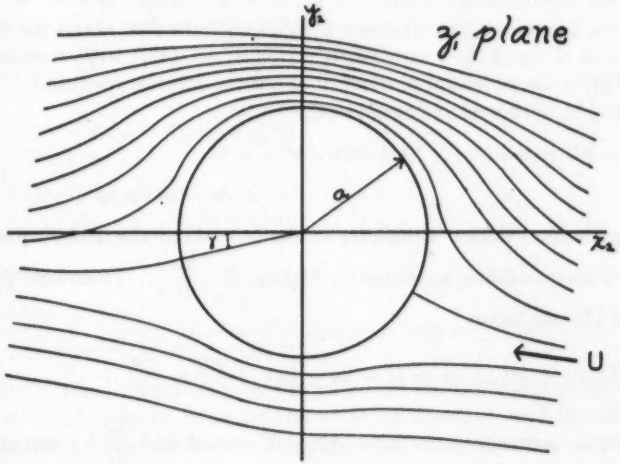


FIG. 7

If K is the value of the circulation, the flow shown above in fig. 6 has the following potential;[9]

$$W = -U\left(z_2 + \frac{a^2}{z_2}\right) - i\frac{K}{2\pi}\ln\frac{z_2}{a}. \qquad (4.1)$$

[9] Glauert, H.; Aerofoil and Airscrew Theory, page 56.

To get a flow making an angle of attack with the axis of reals of α, let

$$z_1 = z_2 e^{-i\alpha} \tag{4.2}$$

and map the flow of fig. 6 upon the z_1 plane of fig. 7 by means of the above conformal mapping function. We have then, in the z_1 plane of fig. 7, a flow making an angle of attack α with the axis of reals, and having a velocity at infinity of magnitude U.

We shall put the flow in the z_1 plane about the arbitrary aerofoil, determining the angle γ of fig. 7 by Joukovsky's hypothesis, that the rear stagnation point in the z_1 plane goes into the tail point of the arbitrary aerofoil (and the base aerofoil as well, as we will always make these two points coincide). The potential of the flow about the circle $| z_1 | = a$ in the z_1 plane of fig. 7 is, substituting eq. (4.2) in eq. (4.1);

$$W = -U\left(z_1 e^{i\alpha} + \frac{a^2 e^{-i\alpha}}{z_1}\right) - i\frac{K}{2\pi}\left(\ln\frac{z_1}{a} + i\alpha\right). \tag{4.3}$$

For the rear stagnation point in the z_1 plane to be at the point $ae^{\pi+\gamma}$, in eq. (4.3) K must have the value[10] $4\pi a\, U \sin(\alpha + \gamma)$.

By our fundamental theorem of §1, $z_1 = ae^{Q}$, where $Q = G' + iH'$, maps the flow about the arbitrary aerofoil onto the flow about the circle $| z_1 | = a$ of fig. 7, and vice versa. Hence, eq. (4.3) with z_1 replaced by ae^{Q} gives the potential of the flow about the arbitrary aerofoil. This potential is, with a as yet undetermined;

$$\begin{aligned} W = -aU\,[&\cos\alpha(e^{Q} + e^{-Q}) + i\sin\alpha(e^{Q} - e^{-Q}) \\ &+ 2\,(iQ - \alpha)\sin(\alpha + \gamma)]. \end{aligned} \tag{4.4}$$

What is the velocity at infinity of the flow about the arbitrary aerofoil? This question is answered by finding $\lim_{\zeta\to\infty}\left|\frac{dw}{d\zeta}\right|$. From eqs. (3.10) and (3.15), we have

$$G'(\rho,\varphi) + iH'(\rho,\varphi) = Q = a_0 + \ln\zeta + \frac{ra_1}{\zeta} + \frac{r^2 a_2}{\zeta^2} + \cdots \tag{4.5}$$

where $\zeta = \rho e^{i\varphi}$ and $a_n = \alpha_n - i\beta_n$. We shall find $\frac{dw}{d\zeta}$ by use of the following implicit differentiation identity;

$$\frac{dw}{d\zeta} = \frac{dw}{dQ} \times \frac{dQ}{d\zeta}. \tag{4.6}$$

[10] Glauert, H.; Aerofoil and Airscrew Theory, page 83.

From eq. (4.4), we have

$$\frac{dw}{dQ} = -aU\left[\cos\alpha(e^{Q} - e^{-Q}) + i\sin(e^{Q} + e^{-Q}) + 2i\sin(\alpha+\gamma)\right]. \quad (4.7)$$

Differentiating eq. (4.5) with respect to ζ, we get

$$\frac{dQ}{d\zeta} = \frac{1}{\zeta} - \frac{ra_1}{\zeta^2} - \frac{2r^2a_2}{\zeta^3} \cdots . \quad (4.8)$$

Using eq. (4.5) again, we find e^{Q} and e^{-Q} to be given by the expressions;

$$\left.\begin{aligned} e^{Q} &= \zeta e^{a_0 + \frac{ra_1}{\zeta} + \frac{r^2a_2}{\zeta^2} \cdots} \\ \text{and} \qquad e^{-Q} &= \frac{1}{\zeta} e^{-a_0 - \frac{ra_1}{\zeta} - \frac{r^2a_2}{\zeta^2} \cdots} \end{aligned}\right\}. \quad (4.9)$$

Putting the values we have found in eqs. (4.7) and (4.8) in eq. (4.6), we have;

$$\begin{aligned} \frac{dw}{d\zeta} = -aU\Bigg[&\cos\alpha\left(\zeta e^{a_0 + \frac{ra_1}{\zeta} + \cdots} - \frac{1}{\zeta} e^{-a_0 - \frac{ra_1}{\zeta} \cdots}\right) \\ &+ i\sin\alpha\left(\zeta e^{a_0 + \frac{ra_1}{\zeta} + \cdots} + \frac{1}{\zeta} e^{-a_0 - \frac{ra_1}{\zeta} \cdots}\right) \\ &+ 2i\sin(\alpha+\gamma)\Bigg]\left[\frac{1}{\zeta} - \frac{ra_1}{\zeta^2} - \frac{2r^2a_2}{\zeta^3} \cdots\right], \end{aligned} \quad (4.10)$$

from which

$$\lim_{\zeta\to\infty}\left|\frac{dw}{d\zeta}\right| = \lim_{\zeta\to\infty}\left|-ae^{a_0}U(\cos\alpha + i\sin\alpha)\right| = ae^{a_0}U. \quad (4.11)$$

In the z_1 plane above and in the plane of the aerofoils of fig. 4, the velocities at infinity must be the same in direction and the same in magnitude. The conformal mapping that exists between these two planes insures the velocities at infinity having the same direction. By the proper choice of the constant a in eq. (4.11), the magnitude of the velocity at infinity in the plane of the aerofoils can be made equal to U, which is the magnitude of the velocity at infinity in the z_1 plane of fig. 7. We make the magnitude of the velocity at infinity in the plane of the

aerofoils be equal to U by making the coefficient of U in eq. (4.11) equal to unity; i.e.,

$$a = e^{-a_0} \tag{4.12}$$

a_0 is the constant term of the series for $G' + iH'$ given by eq. (4.5). Knowing the value of a, the radius of the circle $|z_1| = a$ (which maps into the arbitrary aerofoil) is fixed so that the magnitude of the velocities at infinity in both the z_1 plane and the plane of the aerofoils is U.

As was mentioned earlier in this section, the angle γ in the z_1 plane (fig. 7) is determined by Joukovsky's hypothesis that the tail point of the arbitrary aerofoil must be a stagnation point of the flow about the arbitrary aerofoil. We will therefore find γ by finding what point in the z_1 plane the tail point of the arbitrary aerofoil goes into. By our fundamental theorem of §1, $z_1 = ae^{G'+iH'}$ maps conformally the region exterior to the arbitrary aerofoil on the region exterior to the circle $|z_1| = a$ in the z_1 plane. At the tail point of the arbitrary aerofoil, which is a boundary point of the region exterior to the arbitrary aerofoil, $G' = 0$.[11] Therefore the tail point of the arbitrary aerofoil goes into the point

$$z_1 = ae^{iH'} = ae^{i(H+\delta H)}, \tag{4.13}$$

where H' is the value of the conjugate function for the arbitrary aerofoil at the tail point, H is the value of the conjugate function for the base aerofoil taken at the tail point, and δH is the change in the conjugate function of the base aerofoil taken at the tail point as the base aerofoil changes into the arbitrary aerofoil. δH is calculated by means of eq. (3.6). Noticing the meaning of γ in fig. 7, by what has been said in the discussion preceding eq. (4.13), it is seen at once that γ equals the value at the tail point of the base aerofoil of $H + \delta H - \pi$. That is,

$$\gamma = [H + \delta H]_{\zeta=-2c} - \pi. \tag{4.14}$$

With the constant a determined by eq. (4.12), the quantity Q given by eq. (4.5) and the angle γ determined by eq. (4.14), eq. (4.4) for the potential of the flow about the arbitrary aerofoil becomes;

$$W = -e^{-a_0}U[\cos\alpha(e^Q + e^{-Q}) + i\sin\alpha(e^Q - e^{-Q}) + 2(iQ - \alpha)\sin(\alpha + \gamma)]. \tag{4.15}$$

[11] Kellogg, O. D.; Foundations of Potential Theory, page 364.

This flow has the rear stagnation point at the tail point of the arbitrary aerofoil, a velocity of magnitude U at infinity, and makes an angle of attack α with the chord of the arbitrary aerofoil (the axis of reals).

5. **Calculation of the Lift of the Arbitrary Aerofoil.** If the potential of the flow about an aerofoil is known, the lift of the aerofoil may be determined by means of the following formula;[12]

$$X - iY = \frac{1}{2}\rho i \int_{C'} \left(\frac{dw}{d\zeta}\right)^2 d\zeta , \tag{5.1}$$

where W is the potential of the flow about the aerofoil, C' is a path of integration surrounding the aerofoil, and X and Y are the components, parallel to the ξ and η axes, of the resultant force on the aerofoil. (See fig. 4.)

First, let us work out a series expansion for $\left(\frac{dw}{d\zeta}\right)^2$, on the circle of radius r (C' of fig. 4). Using the well known series expansion for e^z, we write

$$e^Q = 1 + Q + \frac{Q^2}{2!} + \frac{Q^3}{3!} + \cdots .$$

From eq. (4.5), we have

$$Q = G' + iH' = a_0 + \ln \zeta + \frac{ra_1}{\zeta} + \frac{r^2 a_2}{\zeta^2} + \cdots . \tag{5.2}$$

Using the above series in the expansion for e^Q, we find that

$$e^Q\big|_{\zeta = re^{i\theta}} = C_1 e^{i\theta} + C_2 + C_3 e^{-i\theta} + \cdots \tag{5.3}$$

where

$$C_1 = r\left(1 + a_0 + \frac{a_0^2}{2!} + \frac{a_0^3}{3!} + \cdots\right) = re^{a_0},$$

$$C_2 = a_1 C_1$$

and

$$C_3 = \left(\frac{a_1^2}{2!} + a_2\right) C_1 .$$

[12] Glauert, H.; Aerofoil and Airscrew Theory, page 81.

In a similar manner, we find

$$e^{-Q}\big|_{\zeta=re^{i\theta}} = D_3 e^{-i\theta} + \cdots \tag{5.4}$$

where

$$D_3 = \frac{1}{r}\left(1 - a_0 + \frac{a_0^2}{2!} - \frac{a_0^3}{3!} + \cdots\right) = \frac{1}{r}\, e^{-a_0}.$$

From eq. (4.15)

$$\frac{dw}{dQ} = -e^{-a_0} U(E_1 e^{i\theta} + E_2 + E_3 e^{-i\theta} + \cdots), \tag{5.5}$$

where

$$E_1 = C_1 e^{i\alpha},$$

$$E_2 = C_2 e^{i\alpha} + 2i \sin(\alpha + \gamma)$$

and

$$E_3 = C_3 e^{i\alpha} - D_3 e^{-i\alpha}.$$

From eq. (5.2)

$$\left.\frac{dQ}{d\zeta}\right|_{\zeta=re^{i\theta}} = \frac{1}{r}(e^{-i\theta} - a_1 e^{-2i\theta} - 2a_2 e^{-3i\theta} \cdots). \tag{5.6}$$

Multiplying eqs. (5.5) and (5.6) together, we get

$$\frac{dw}{d\zeta} = -\frac{e^{-a_0}}{r} U(F_0 + F_1 e^{-i\theta} + F_2 e^{-2i\theta} + \cdots) \tag{5.7}$$

where

$$F_0 = E_1,$$

$$F_1 = E_2 - a_1 E_1 = 2\, i \sin(\alpha + \gamma)$$

and

$$F_2 = E_3 - a_1 E_2 - 2\, a_2 E_1.$$

Squaring eq. (5.7), we find

$$\left(\frac{dw}{d\zeta}\right)^2_{\zeta=re^{i\theta}} = \frac{e^{-2a_0}}{r^2} U^2(K_0 + K_1 e^{-i\theta} + K_2 e^{-2i\theta} + \cdots), \tag{5.8}$$

where

$$K_0 = F_0^2,$$

$$K_1 = 2F_0F_1,$$

and

$$K_2 = 2F_0F_2 + F_1^2.$$

Later we shall see that the value of the lift depends on the constant K_1 and that the value of the moment depends upon the real part of the constant iK_2. It will be found convenient to have the constants K_1 and iK_2 expressed in terms of quantities that are as near fundamental as possible. The following formulae, which are calculated from the preceding relations, give a rapid way of computing K_1 and "R"(iK_2), where "R"(iK_2) means real part of iK_2.

$$K_1 = 4iC_1 e^{i\alpha} \sin(\alpha + \gamma) \tag{5.9}$$

$$\text{"}R\text{"}(iK_2) = 2C_1 \text{"}R\text{"}[2a_1 e^{i\alpha} \sin(\alpha + \gamma) - iC_2 e^{2i\alpha}]. \tag{5.10}$$

Applying eq. (5.1), and taking the path of integration to be the circle C' of radius r in fig. 4, we find

$$X - iY = \frac{1}{2}\rho i \int_0^{2\pi} \frac{e^{-2a_0}}{r^2} U^2 (K_0 + K_1 e^{-i\theta} + \cdots) ire^{i\theta} d\theta = -\rho\pi \frac{e^{-2a_0}}{r} K_1 U^2.$$

Substituting the value of K_1 given in eq. (5.9) in the above equation, we have

$$X - iY = -4i\rho\pi e^{-a_0} e^{i\alpha} U^2 \sin(\alpha + \gamma). \tag{5.11}$$

Equating real parts and imaginary parts in the above relation, we have

$$X = 4\rho\pi e^{-a_0} U^2 \sin(\alpha + \gamma) \sin\alpha, \tag{5.12}$$

and

$$Y = 4\rho\pi e^{-a_0} U^2 \sin(\alpha + \gamma) \cos\alpha. \tag{5.13}$$

Eqs. (5.12) and (5.13) show that X and Y are components of a force which is perpendicular to the wind direction. This force is the lifting force given by the equation

$$L = \sqrt{X^2 + Y^2} = 4\rho\pi e^{-a_0} U^2 \sin(\alpha + \gamma), \tag{5.14}$$

where a_0 is the constant term of the series (5.2), and γ is determined by eq. (4.14). Instead of writing a lift coefficient in general terms from eq. (5.14), we shall compute the lift of an arbitrary aerofoil using eq. (5.14) and then, if desired, write down the lift coefficient for that particular aerofoil.

6. Calculation of the Moment of the Arbitrary Aerofoil about Its

Leading Edge. The following formula[13] gives the moment M_0 of the arbitrary aerofoil about the origin of coördinates;

$$M_0 = -\frac{1}{2}\rho "R" \int_{C'} \left(\frac{dw}{d\zeta}\right)^2 \zeta \, d\zeta \,, \tag{6.1}$$

where "R" means the real part of the integral. Using the value of $\left(\frac{dw}{d\zeta}\right)^2$ given by eq. (5.8), and using the circle C' of radius r of fig. 4, as a path of integration, we find

$$\begin{aligned} M_0 &= -\frac{1}{2}\rho "R" \int_0^{2\pi} \frac{e^{-2\alpha_0}}{r^2} U^2(\cdots + K_2 e^{-2i\theta} + \cdots) ir^2 e^{2i\theta} \, d\theta \\ &= -\rho\pi e^{-2\alpha_0} U^2 "R"(iK_2) \end{aligned} \tag{6.2}$$

where "R" (iK_2) is calculated by means of eq. (5.10). Having the value of M_0, the moment M about the leading edge of the arbitrary aerofoil is given approximately by the relation

$$M = M_0 - L \text{ (distance from the origin to the leading edge).} \tag{6.3}$$

As in the case of the lift, a moment coefficient (K_M) for the moment about the leading edge of the arbitrary aerofoil will be written down, if desired, after the moment M has been computed by means of eq. (6.3).

PART II

In this part of the paper, the method developed in the first part will be applied to a concrete case. We shall take an arbitrary aerofoil, select a nearby Joukovsky aerofoil, and actually compute the lift and the moment of the arbitrary aerofoil. In order to have a check on the accuracy of the work, we shall take as the arbitrary aerofoil a Joukovsky aerofoil. Let the arbitrary aerofoil be the Joukovsky aerofoil given by table #9 of Reports and Memoranda #1106. In figs. 1 and 2, this aerofoil is given by letting $a = 1.15$, $c = 1$ and $\beta = 6°$. Select as the nearby Joukovsky aerofoil (the base aerofoil) the Joukovsky aerofoil given in table #7 of the same report. In figs. 1 and 2 this aerofoil is given by letting $a = 1.15$, $c = 1$ and $\beta = 0°$. Joukovsky's transformation in each of the above cases takes the form $\zeta = z + \frac{1}{z}$. The table below gives the coördinates of the points on the two aerofoils.

[13] Glauert, H.; Aerofoil and Airscrew Theory, page 81.

TABLE I

Arbitrary aerofoil		Base aerofoil	
Upper surface	Lower surface	Upper surface	Lower surface
+2.065, +.008	+2.062, −.016	+2.069, +.000	Reflection of upper surface in the axis of reals
+2.060, +.049	+2.050, −.046	+2.063, +.041	
+2.043, +.092	+2.030, −.074	+2.053, +.065	
+2.027, +.118	+2.002, −.100	+2.043, +.083	
+1.998, +.153	+1.966, −.123	+2.012, +.120	
+1.962, +.188	+1.869, −.161	+1.814, +.236	
+1.919, +.222	+1.740, −.187	+1.661, +.287	
+1.779, +.305	+1.642, −.197	+1.569, +.301	
+1.694, +.343	+1.533, −.201	+1.458, +.318	
+1.598, +.379	+1.412, −.200	+1.334, +.332	
+1.493, +.411	+1.279, −.193	+1.203, +.339	
+1.378, +.439	+1.135, −.182	+1.063, +.343	
+1.256, +.463	+0.981, −.167	+0.914, +.343	
+1.126, +.483	+0.817, −.148	+0.760, +.336	
+0.990, +.497	+0.649, −.128	+0.603, +.329	
+0.847, +.507	+0.464, −.102	+0.432, +.313	
+0.699, +.511	+0.277, −.077	+0.262, +.296	
+0.546, +.513	+0.085, −.052	+0.088, +.275	
+0.390, +.505	−0.117, −.025	−0.094, +.253	
+0.232, +.493	−0.307, −.005	−0.265, +.226	
+0.071, +.477	−0.503, +.015	−0.441, +.201	
−0.097, +.457	−0.696, +.031	−0.616, +.173	
−0.254, +.431	−0.884, +.043	−0.788, +.146	
−0.416, +.402	−1.064, +.051	−0.954, +.121	
−0.576, +.370	−1.234, +.054	−1.114, +.096	
−0.733, +.336	−1.393, +.056	−1.265, +.075	
−0.886, +.300	−1.533, +.050	−1.407, +.055	
−1.034, +.263	−1.658, +.042	−1.537, +.039	
−1.177, +.226	−1.766, +.033	−1.655, +.025	
−1.312, +.189	−1.853, +.023	−1.756, +.015	
−1.440, +.154	−1.921, +.013	−1.842, +.008	
−1.556, +.121	−1.967, +.006	−1.910, +.004	
−1.664, +.090	−1.994, +.001	−1.959, +.001	
−1.758, +.064	−2.000, .000	−2.000, .000	
−1.839, +.041			
−1.902, +.025			
−1.954, +.011			
−1.987, +.003			

Draw these two aerofoils on a scale that allows a distance equal to 5 percent of the chord to be laid off with about 1 percent accuracy. These two aerofoils are drawn on a small scale in fig. 4. Measure

TABLE II

θ	ξ	δN	M	$\delta N_1 = \frac{\delta N}{M}$
0°	+2.070	−.006	0.41	−.013
10°	+2.044	+.003	0.46	+.007
20°	+1.968	+.019	0.62	+.032
30°	+1.842	+.042	0.80	+.052
40°	+1.671	+.068	1.00	+.068
50°	+1.459	+.099	1.18	+.084
60°	+1.204	+.131	1.36	+.097
70°	+0.913	+.160	1.50	+.107
80°	+0.602	+.184	1.64	+.112
90°	+0.263	+.198	1.73	+.114
100°	−0.089	+.202	1.78	+.113
110°	−0.440	+.192	1.77	+.109
120°	−0.790	+.172	1.70	+.101
130°	−1.118	+.141	1.56	+.090
140°	−1.408	+.104	1.38	+.075
150°	−1.658	+.064	1.11	+.058
160°	−1.847	+.032	0.77	+.042
170°	−1.964	+.008	0.40	+.021
180°	−2.000	.000	0.00	.000
190°	−1.964	−.009	0.40	−.023
200°	−1.847	−.033	0.77	−.043
210°	−1.658	−.067	1.11	−.061
220°	−1.408	−.112	1.38	−.081
230°	−1.118	−.148	1.56	−.095
240°	−0.790	−.183	1.70	−.107
250°	−0.440	−.210	1.77	−.119
260°	−0.089	−.222	1.78	−.125
270°	+0.263	−.218	1.73	−.126
280°	+0.602	−.204	1.64	−.124
290°	+0.913	−.181	1.50	−.120
300°	+1.204	−.152	1.36	−.112
310°	+1.459	−.119	1.18	−.101
320°	+1.671	−.086	1.00	−.087
330°	+1.842	−.056	0.80	−.070
340°	+1.968	−.031	0.62	−.050
350°	+2.044	−.014	0.46	−.029
360°	+2.070	−.006	0.41	−.013

the normal distance from the base aerofoil to the arbitrary aerofoil along the normals to the base aerofoil. These normals were drawn by the "mirror" method. The work from here on follows very closely the theoretical development of the first part of this paper. The symbols

in table II refer to figs. 4 and 5, first part. The column headed θ is the angle of points on the circle C_3 (fig. 5); that headed ξ gives the ξ coördinate of the corresponding point on the aerofoil B (fig. 4); that headed δN gives the normal distances in the aerofoil plane from the base aerofoil to the arbitrary aerofoil; that headed δN_1 gives the normal distances in the z_1 plane (fig. 5) of the curve C_2 from the circle C_3. M is the magnification factor used in going from the z_1 plane to the ζ plane, and is equal to the absolute value of $d\zeta$ on the circle C_3 (fig. 5). For the case we are considering, the circle C_3 is

$$| z_1 | = 1.15 \text{ and } \zeta = \xi + i\eta = z_1 + .15 + \frac{1}{z_1 + .15}.$$

Now plot δN_1 vs. θ and get the Fourier series for δN_1. A Coradi Harmonic Analyzer was used in getting all Fourier coefficients. The Fourier series for δN_1 is (eq. 3.1, first part);

$$\delta N_1 = -.006 - .004 \cos\theta + .120 \sin\theta - .001 \sin 2\theta - .001 \cos 3\theta - .001 \cos 4\theta \cdots.$$

In fig. 4, take the circle C' as $| \zeta | = 2.5$. We will now find what points (P_1) of fig. 5, the points P of fig. 4 go into. Using eqs. (2.8) and (2.9) and the relation $z_1 = z - .15$, we get the P_1 points corresponding to the P points on the circle $| \zeta | = 2.5$. Using eq. (3.5), we calculate $\delta G_1(P_1)$, which is equal to $\delta G(P)$. On the circle $| \zeta | = 2.5$ (C', fig. 4), take the P points at 10° intervals. Table III gives the resulting values of $\delta G(P)$.

Now plot $\delta G(P)$ against ψ and get the Fourier series given by eq. (3.7), from which get the Fourier series given by eq. (3.8). The constant term and two pairs of coefficients are sufficient in these series. Eq. (3.8) for our arbitrary aerofoil is;

$$\delta G(\rho, \varphi) = +.005 + .003 \frac{\cos\varphi}{\rho} - .120 \frac{\sin\varphi}{\rho} + .001 \frac{\cos 2\varphi}{\rho^2} - .017 \frac{\sin 2\varphi}{\rho^2} \cdots.$$

Eq. (3.9) for our base aerofoil is;

$$G(\rho, \varphi) = -.140 + \ln\rho - .150 \frac{\cos\varphi}{\rho} - 1.011 \frac{\cos 2\varphi}{\rho^2} \cdots.$$

TABLE III

ψ	$P(\xi,\eta)$	$P_1(r_1,\psi_1)$	$\delta G(P)$
0°	+2.50, +0.000	+1.85, + 0°.0	+.0074
10°	+2.46, +0.434	+1.93, + 17°.1	−.0110
20°	+2.35, +0.855	+2.10, + 30°.9	−.0209
30°	+2.17, +1.25	+2.27, + 42°.0	−.0289
40°	+1.92, +1.61	+2.44, + 51°.6	−.0326
50°	+1.61, +1.92	+2.57, + 60°.7	−.0352
60°	+1.25, +2.17	+2.67, + 69°.2	−.0367
70°	+0.855, +2.35	+2.76, + 77°.3	−.0372
80°	+0.434, +2.46	+2.82, + 85°.2	−.0374
90°	0.000, +2.50	+2.85, + 93°.0	−.0373
100°	−0.434, +2.46	+2.86, +101°.	−.0366
110°	−0.855, +2.35	+2.84, +109°.	−.0356
120°	−1.25, +2.17	+2.80, +116°.	−.0342
130°	−1.61, +1.92	+2.73, +125°.	−.0313
140°	−1.92, +1.61	+2.63, +133°.	−.0292
150°	−2.17, +1.25	+2.50, +143°.	−.0255
160°	−2.35, +0.855	+2.37, +153°.	−.0196
170°	−2.46, +0.434	+2.22, +165°.	−.0108
180°	−2.50, 0.000	+2.15, +180°.	+.0030
190°	−2.46, −0.434	+2.22, −165°.	+.0170
200°	−2.35, −0.855	+2.37, −153°.	+.0265
210°	−2.17, −1.25	+2.50, −143°.	+.0329
220°	−1.92, −1.61	+2.63, −133°.	+.0373
230°	−1.61, −1.92	+2.73, −125°.	+.0397
240°	−1.25, −2.17	+2.80, −116°.	+.0429
250°	−0.855, −2.35	+2.84, −109°.	+.0448
260°	−0.434, −2.46	+2.86, −101°.	+.0461
270°	0.000, −2.50	+2.85, − 93°.0	+.0573
280°	+0.434, −2.46	+2.82, − 85°.2	+.0476
290°	+0.855, −2.35	+2.76, − 77°.3	+.0478
300°	+1.25, −2.17	+2.67, − 69°.2	+.0477
310°	+1.61, −1.92	+2.57, − 60°.7	+.0466
320°	+1.92, −1.61	+2.44, − 51°.6	+.0446
330°	+2.17, −1.25	+2.27, − 42°.0	+.0414
340°	+2.35, −0.855	+2.10, − 30°.9	+.0369
350°	+2.46, −0.434	+1.93, − 17°.1	+.0252
360°	+2.50, 0.000	+1.85, 0°.0	+.0074

Combining the two above series, eq. (3.10), for our case, becomes

$$G'(\rho,\varphi) = -.135 + \ln\rho - .147\frac{\cos\varphi}{\rho} - .120\frac{\sin\varphi}{\rho} - 1.010\frac{\cos 2\varphi}{\rho^2}$$

$$- .017\frac{\sin 2\varphi}{\rho^2} \cdots$$

and eq. (3.15) is

$$H'(\rho, \varphi) = \varphi - .120\frac{\cos\varphi}{\rho} + .147\frac{\sin\varphi}{\rho} - .017\frac{\cos 2\varphi}{\rho^2} + 1.010\frac{\sin 2\varphi}{\rho^2} \cdots .$$

Eq. (4.5) becomes, combining the two series above,

$$G' + iH' = -.135 + \ln\zeta - \frac{.147 + .120i}{\zeta} - \frac{1.010 + .017i}{\zeta^2} \cdots$$

i.e., $a_0 = -.135$, $a_1 = -.059 - .048\,i$ and $a_2 = -.162 - .003\,i$.

Using eq. (3.6), we find $\delta H_1(P_1)$ when P is the tail point of the aerofoils. We thus get $\delta H(P)$, when P is the tail point of the aerofoils, by findings its equal $\delta H_1(P_1)$. In our case the tail point is $\zeta = -2$, and the corresponding P_1 point is $z_1 = -1$. Eq. (3.6) gives $\delta H_1(+1, \pi) = 6°0'20''$. From eq. (4.14) $\gamma = \delta H(+2, \pi) = 6°0'20'' = .105$ radians. This should come out exactly 6°, so we are in error only a small fraction of a percent.

From eq. (5.14), the lift of our arbitrary aerofoil is;

$$L = 4.58\,\pi\,\rho\,U^2 \sin(\alpha + .105).$$

Glauert[14] gives as the lift of the Joukovsky aerofoil we are using for the arbitrary aerofoil;

$$L = 4.60\,\pi\,\rho\,U^2 \sin(\alpha + .105),$$

so we have a deviation of only a fraction of a percent in our determination of the lift of the arbitrary aerofoil. Using that part of the axis of reals intercepted by the arbitrary aerofoil as the chord (= 4.06 in our case), we find that the lift coefficient of our arbitrary aerofoil is nearly;

$$K_L = 3.52\,(\alpha + .105).$$

Let us next compute the moment of our arbitrary aerofoil about its leading edge. We first calculate M_0 using eq. (6.2);

$$M_0 = -\rho\pi U^2 e^{-2a_0}\,\text{“}R''(iK_2),$$

where “$R''(iK_2)$ is given by eq. (5.10). Using this equation, we find “$R''(iK_2) = -3.55\alpha - .054$ and eq. (6.2) gives us

$$M_0 = \rho\pi U^2[4.65\alpha + .071].$$

[14] Glauert, H.; Aerofoil and Airscrew Theory, page 85.

Glauert[14] gives as the value of M_0 for the Joukovsky aerofoil we are using as the arbitrary aerofoil, the following;

$$M_0 = \rho\pi U^2[4.60\alpha + .071],$$

so we have determined M_0 to within about 1 percent of what a rigorous theoretical calculation gives. Making the approximation that the sine of a small angle equals the angle, eq. (6.3) gives the following value for the moment of our arbitrary aerofoil about its leading edge;

$$M = -\rho\pi U^2[4.65\alpha + .907]$$

from which we have

$$K_m = -.886\alpha - .172 = -.251\, K_L - .089.$$

In this numerical case, which has been carried through, the maximum normal distance from the base aerofoil to the arbitrary aerofoil was about 5½ percent of the chord, which is quite large. Consequently, one base aerofoil may serve as the base aerofoil for a large number of arbitrary aerofoils. A dozen properly selected base aerofoils (all Joukovsky's with equal chords) would be a sufficient number of base aerofoils to investigate a great number of arbitrary aerofoils. That is to say, a large number of arbitrary aerofoils (drawn with the proper chord) would be within the above distance of members of the base family. Certain constants are calculated once and for all for each base aerofoil; e.g., in tables II and III, the columns headed P, P_1, ξ and M are such constants. Thus, if we are given an arbitrary aerofoil, we can, with a few hour's work, find the lift and the moment characteristics of the arbitrary aerofoil.

SOME PROBLEMS OF MATHEMATICAL BIOLOGY

By J. B. S. Haldane[1]

It appears to be the fate of all sciences to become more or less completely permeated by mathematics. In the case of biology, the time when it will resemble modern physics is extremely remote, and a biologist with mathematical leanings like myself must at once testify that some of his ablest colleagues who are advancing our science most rapidly are completely incapable of performing the most elementary mathematical operations, and are none the worse as biologists for this fact.

Biological problems occur, I think, at three levels. The most fundamental level is the analysis of the life of the cell. The second is the analysis of a whole organism such as a man or an oak, taking the cells as units. The third level is concerned with the relations of a number of organisms to one another, for example members of a population of one species, or parasites and hosts.

On the first level, that of cell life, we have a few problems arising out of enzyme chemistry, and some problems of diffusion of heat and solutes which have been very elegantly solved by A. V. Hill. But these are essentially physical and chemical problems, not requiring really new technique. However if, as Dirac believes, there is ever to be a specific mathematics of life, it will be required at this level. Just as mathematical chemistry, the science whose embryonic outlines have been blocked out in the last ten years, is much more complicated than the old mathematical physics, so mathematical biology will no doubt require very great subtlety to express the wholeness of living beings.

Nevertheless it is worth noting how, even at its present level, wave mechanics bridges the gap between chemical and biological notions. Consider the reproduction of a cell. Every gene is represented by two like itself, one in each daughter cell. The gene is about the size of a protein molecule, highly specific, and cannot possibly divide into two equal halves, like a drop of water. The chemist says one of the daughter genes must be the model, the other the copy. The biologist brings in

[1] Professor of Genetics at the University of London. This paper was delivered as the Sedgwick Memorial lecture of the Biology Department of the Massachusetts Institute of Technology.—*Ed.*

the notion of heredity, and regards both as heirs. The mathematician can point out that there is no way of distinguishing between the two alternatives. Just where the vitalist and mechanist are most clearly at grips, the mathematician can order them to break away. But he cannot yet offer an adequate substitute for their views.

On the second biological level, the level at which the relations of cells are discussed, L. J. Henderson has shown that nomography offers peculiar advantages in the analysis of the very complicated equilibria which occur. On its way round the circulation the blood absorbs and gives up O_2 and CO_2. This gives us two degrees of freedom. But we have a very large number of variables, for example amount and pressure of the two gases, percent saturation of the haemoglobin, osmotic pressure, cell volume, and so on. It is possible to represent the relations of all these variables, not by a surface in twelve or more dimensions, but by a nomogram in two only.

One important branch of physiology to-day is the study of the hormones. A hormone is a substance produced in one organ and regulating the activity of another. Two hormones may have a similar effect on the same function. For example adrenaline and thyroxine both increase basal metabolism. Or they may have opposite effects, for example insulin tends to lower blood sugar, adrenaline to raise it. The hormone concentration in the blood and the activities controlled by them would thus form a system with some thousands of variables but only perhaps about twenty degrees of freedom. Will it be possible to represent the whole system on a twenty dimensional nomogram? If so will it be possible to read off the value of one variable in terms of the others by a suitable apparatus, as we now solve systems of simultaneous equations electrically?

If the answer is yes, the future holds in store a mathematical medicine which will solve problems that to-day are solved more or less intuitively, just as the mediaeval architects solved (or sometimes failed to solve) mechanical problems which are to-day matters of exact calculation.

But I must turn to the third level of biology, the level of populations. Here there are two classes of problem, those of distribution and of estimation. In the first class we ask, given certain laws, what sort of population do we expect, and how should it change. In the second class we have a sample, and must inquire what sort of population it represents. For example we might first calculate the effects of inbreeding on a population, and then catch a hundred flies, and see whether the statistical distribution of characters among them differs significantly from that

found in a population mating at random, i.e. whether there appears to be inbreeding in this case.

Suppose we have a population composed of n different sorts of individuals, say black, white, and gray rabbits, or healthy men, sick men, and hookworms, we naturally represent it by a point in a space of n dimensions, but if we are only concerned with the proportions of the different kinds we represent it in a simplex of $n - 1$ dimensions. For example we can represent our three sorts of rabbit, if we are not concerned with the size of the rabbit population, by a point in a triangle.

Once we have adopted this method of representation we can state a number of our problems geometrically. For example what will be the locus of all points representing populations of rabbits mating at random? What will be the course taken by such a representative point if the breeding system is altered, or if one kind of rabbit is favored by natural selection at the expense of others?

Let us consider first the population in which the variation due to one pair of genes only is considered. I remind you that a gene is a unit in a nucleus such that only one gene of a kind can get into a gamete, and only two into a zygote. If the two of a kind in a zygote are completely alike, it is called a homozygote, if they differ to some extent, a heterozygote, which produces equal numbers of gametes carrying each of the two slightly different sorts of gene. For example if we call our genes A and a, then in the guineapig AA individuals could be yellow, Aa cream, aa white. But more usually AA and Aa are indistinguishable, a condition known as dominance. Thus in man AA and Aa might be normal in appearance, and aa albinos.

Now suppose we call the proportions of AA, Aa, aa in the population x, y, and z, where $x + y + z = 1$, it is easy to show that in a population mating at random $y^2 = 4\,xz$. The locus of the representative points of such populations is a parabola whose vertex is at the center of one altitude of the equilateral triangle X Y Z, and touching the sides XY and ZY at X and Z. Supposing we subject such a population to natural selection, preserving the system of random mating, its representative point should move along this parabola. Supposing we alter the mating system, by introducing say universal brother-sister mating, the representative point will move down along the perpendicular PQ in Fig. 1.

Any gene frequency is representated by a perpendicular line on this diagram, and any mating system leads to an equilibrium either with $y = 0$ or represented by a parabola,

$$4xz - y^2 = \frac{2\alpha y}{1 - \alpha},$$

where α is the mean value of a quantity defined by Sewall Wright, and called the coefficient of inbreeding. In human populations α is generally less than .01.

Similarly if we take a slightly more complex case, that of the human blood groups, everyone belongs to group O, A, B, or AB. If x, y, z, t, are the proportions of a population belonging to these groups, then Bernstein showed that in a random mating population

$$(x + y)^{\frac{1}{2}} + (x + z)^{\frac{1}{2}} = x^{\frac{1}{2}} + (x + y + z + t)^{\frac{1}{2}}$$

If we represent our population by a point in a tetrahedron, this is the equation of a quartic surface. Bernstein found that the representative

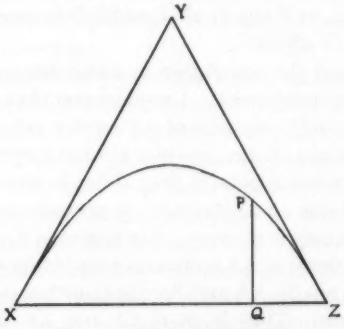


FIG. 1. (After de Finetti)

points for populations which had been well mixed in the past lay near this surface. But a population consisting of different races which had never mixed, for example the hospital patients in the port of Liverpool, gave a point lying well off the surface.

In general a mating system or a system of selection may do one of two things. It may lead to a degenerate equilibrium in which certain types of organism have disappeared, for example continued mating of double first cousins leads to a system with no heterozygotes. Or it may lead to a non-degenerate system represented by a locus in the representative space. Wright showed that this was the case with continued mating of single first cousins. If we allow for mutation, the process by which new genes arise, no equilibria are quite degenerate.

These statical problems are very interesting where we have a polymorphic population of animals, for example the snails of England, and

many species of grasshopper, both American and English. We know that in our snails the presence of bands is recessive to their absence. We also know that the proportion of banded snails has not altered much in the last three thousand years. In this case it looks as if the two kinds of snail were on the whole equally fit in the Darwinian sense.

But in the American grasshoppers, Nabours' work suggests that the equilibrium is stable. If we consider a gene A causing a particular color pattern not found in the commonest type aa, the relative viabilities of the types seem to be in the order Aa > aa > AA. We may regard our representative point as acted on by forces determined by this selective system, but constrained by the mating system to lie on a particular curve. It is then easy to determine the point of equilibrium. Nevertheless in a system involving many gene pairs it is far from simple to determine whether a given equilibrium is stable or not, and a good deal of fallacious verbal reasoning about natural selection and eugenics has been based on the tacit assumption that equilibria are stable when they are not so in fact.

To turn to dynamical problems, let us first consider populations where the generations do not overlap, for example annual plants such as wheat or poppies, or animals such as many Lepidoptera with exactly one or two annual broods per year. Then we can show that under a given mating system the changes in the population are expressed by a series of simultaneous linear finite difference or recurrence equations. For example, to take a trivial case, consider a population in which differences are due to one pair of genes, and in which brother-sister mating is instituted. It breaks up into groups of two in each generation. In the nth generation let these groups be in the proportions:—

$$p_n\begin{cases}\mathrm{AA}\times\mathrm{AA}\\ \mathrm{aa}\times\mathrm{aa}\end{cases}, \qquad q_n\,\mathrm{AA}\times\mathrm{aa}, \qquad r_n\begin{cases}\mathrm{AA}\times\mathrm{Aa}\\ \mathrm{aa}\times\mathrm{Aa}\end{cases}, \qquad s_n\,\mathrm{Aa}\times\mathrm{Aa}.$$

Then in the next generation:

$$\begin{aligned} p_{n+1} &= p_n + \tfrac{1}{4}r_n + \tfrac{1}{8}s_n \\ q_{n+1} &= \tfrac{1}{8}s_n \\ r_{n+1} &= \tfrac{1}{2}r_n + \tfrac{1}{2}s_n \\ s_{n+1} &= q_n + \tfrac{1}{4}r_n + \tfrac{1}{4}s_n \end{aligned}$$

where $p_n + q_n + r_n + s_n = 1$. The last three equations form an independent system, whose solutions are of the form q_n, r_n, or $s_n = a_1\lambda_1^n + a_2\lambda_2^n + a_3\lambda_3^n$, where $\lambda_1, \lambda_2, \lambda_3$ are the roots of

$$\begin{vmatrix} -\lambda & 0 & \frac{1}{8} \\ 0 & \frac{1}{2} - \lambda & \frac{1}{2} \\ 1 & \frac{1}{4} & \frac{1}{4} - \lambda \end{vmatrix} = 0,$$

i.e. $\lambda = \frac{1}{4}$, or $\frac{1 \pm \sqrt{5}}{4}$. The values of a_1, a_2, a_3, depend on the initial conditions. Wright deals with the same type of problem by a different method which leads to the equation

$$x_{n+2} - \tfrac{1}{2}\, x_{n+1} - \tfrac{1}{4}\, x_n = 0$$

where x_n is the proportion of Aa individuals in the nth generation.

That is of course a trivial problem, but there is nothing in the least trivial about a problem involving linkage between genes which leads to a set of 16 simultaneous linear recurrence equations with algebraic, not numerical, coefficients. I may add that the results of this sort of analysis are well borne out by experiment.

When the population changes, not under the influence of inbreeding or outbreeding, but of selection or mutation, we get a series of non-linear finite difference equations. The simplest possible case is as follows. Consider a random mating population in which the two genes are found in the proportion 1 A: x_n a. The three zygotic types occur in the proportions:

$$1 \text{ AA}: 2\, x_n \text{ Aa}: x_n^2 \text{ aa}$$

Suppose the relative fitness of AA and Aa to be unity, while that of aa is $1 - s$, where s is called the coefficient of selection. Then the parents of the next generation occur in the proportions 1 AA: $2\, x_n$ Aa: $(1 - s)\, x_n^2$ aa.

Hence
$$x_{n+1} = \frac{(1 - s)x_n^2 + x_n}{x_n + 1},$$

or

$$\Delta x_n = \frac{-sx_n^2}{x_n + 1}$$

a non-linear finite difference equation. If s is small, we can of course treat it as a differential equation

$$\frac{dx}{dn} = -\frac{sx^2}{x+1}$$

whence $sn = \frac{1}{x_n} - \log x_n + C$. But in general we are up against the problem of solving the difference equation

$$\Delta x_n = s\,\phi\,(x_n)$$

I have shown that in this case n can be expanded in a series of ascending powers of s. Unfortunately I have not been able to discover when this series converges, though luckily it turns out to be one of those kindly series which are quite useful even when they diverge. But I have failed ignominiously with only slightly more complicated problems. For example I cannot answer such a simple question as this. "The absence of horns in cattle is due to a dominant gene. In a herd of cattle mating at random a proportion p have horns. If all cows, but no horned bulls, are bred from, what will be the proportion with horns after n generations?" Of course numerical solutions can be given, but I put it to mathematicians that non-linear difference equations are a perfectly respectable subject for research, even if they do not yet often occur in physical problems.

When generations overlap, as in man, these difference equations become integral equations, non-linear when they are concerned with selection.

Besides the mating system and the selective system, populations change under the influence of mutation and mere random survival. The latter problem has been specially treated by Wright and Fisher. As a result of fluctuation, that is to say what from the biological view is mere chance, our representative point is replaced, after one generation, by a continuous distribution of points around it, with varying degrees of probability, and the problem can be treated by the methods used for the study of heat diffusion. But near the edge of our diagram, when there are only a small number of individuals carrying a given gene, these methods will not work, any more than they will work in physics where just a few quanta are concerned. In this case the biologist is confronted with another set of non-linear difference equations.

It will be seen that these problems of changes in a population, which

Darwin and his contemporaries tried to solve verbally, are really far from simple. Our theory still runs far ahead of experiment or observation, nevertheless something is being done on these lines. For example last year my colleague Mr. Gordon released a mixed population of 36,000 fruit flies in an English orchard, and followed the changes in it through the year.

But they have importance in another way. Let us suppose that these investigations have been of merely theoretical interest because, say, natural selection and mutation have played no serious part in evolution; yet they have compelled us, for the first time, to define really accurately the words such as "fitness" which Darwin used. Just as our ancestors used the same word "force" to mean what we now call force, torque, energy, momentum, or cause, so the variety of meanings given to this word "fitness" has made accurate discourse on evolution or eugenics rather difficult. We can now see that a number of events which, verbally stated, are paradoxical, may really happen. For example, as a result of the survival of the fittest, a species may become progressively less fit, just as a spinning top may rise up under the influence of gravitation. Whether or not Fisher, Wright, and I have made the Darwinian theory more probable, we have certainly made it more precise. Perhaps we have even opened the way for its disproof, since in some cases it is now possible to calculate the intensities of selection needed to cause an evolutionary change, and these may turn out to be higher than are found in nature.

A wholly different type of problem was investigated by Lotka in this country, Volterra in Italy, Nicholson and Bailey in Australia. They concerned themselves with the absolute number, either of a single species, or two or more species living in a biological association such as competition or parasitism.

I will give two examples of the kind of analysis arising in such cases. Supposing we have two species, say a host and a parasite, whose numbers are $X + x$, and $Y + y$, where X and Y are the values characteristic of a stationary state, that is to say an equilibrium, stable or unstable. Then

$$\frac{dx}{dt} = f(x, y) = ax + by + fxy + \cdots$$

$$\frac{dy}{dt} = F(x, y) = cx + dy + exy + \cdots$$

and in the immediate neighbourhood of the stationary point we need only consider the first two terms of the expansion. Then

$$x = A_1 e^{\lambda_1 t} + A_2 e^{\lambda_2 t} + \cdots$$
$$y = B_1 e^{\lambda_1 t} + B_2 e^{\lambda_2 t} + \cdots$$

where λ_1, λ_2, are the roots of

$$\begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = 0.$$

The treatment is just the same when there are more than two species. Suppose $X + x$ is the number of hosts, $Y + y$ of parasites, then it is clear that in general a, c, and d are generally positive, b negative. Hence the real parts of the roots of

$$\lambda^2 - (a + d)\lambda + ad - bc = 0$$

are positive. So any oscillations round the equilibrium are unstable. The numbers are bound to oscillate. Exact calculation of the periodicity of epidemic waves will naturally involve other terms in the equations. This is of course a highly simplified system. But Nicholson and Bailey in particular have dealt with much more concrete cases, and Gause has recently experimented on these lines with two species of Protozoa. As might be expected, practice was more complicated than theory. Nevertheless theory was a stimulus to experiment, and predictions were at least roughly confirmed.

In the above calculation no account was taken of the life cycle. Let us see what happens if we do this. Consider a human population in which

$N(t)$ = number at time t
$B(t)$ = births per year at time t
$S(x)$ = chance of surviving to age x (or older)
$F(x)$ = chance of producing a child at age x

For simplicity we only refer to females.

Then
$$N(t) = \int_0^\infty B(t - x)\, S(x)\, dx$$
$$B(t) = \int_0^\infty B(t - x)\, S(x) F(x)\, dx$$

This last is a linear integral equation giving the births in any year in terms of past births. It was shown by Lotka that any oscillations due to irregularities in the population are ultimately damped out, and the population finally increases or decreases exponentially. It will increase or decrease according as

$$\int_0^\infty S(x)\, F(x)\, dx$$

is greater or less than unity, in other words, according as 100 girl babies subjected to existing death and birth rates will have more or less than 100 daughters. In most civilized countries, including the United States, this quantity is less than unity. Hence although $N(t)$ is generally increasing, it will certainly fall unless the birth-rates at given ages can be substantially increased.

It is worth noting that the fundamental problem of predicting future population trends has only quite recently been solved, mainly by Dublin and Lotka, as the result of analysis on these lines.

I now turn to the second type of problem, that of deductions to be made from samples, in other words, to problems of inverse probability. I take two examples from the problem of linkage. Two genes are said to be linked when they are located in the same chromosome. As the result of this, if they enter an organism from the same parent they tend to go out of it into the same child. Thus in mice we have the recessive genes for whiteness and head-shaking in the same chromosome. A mouse whose parents were normal and white shaker is itself normal. If we mate it to a white shaker, 97% of the offspring are normals and white shakers, only 3% are white non-shakers and colored shakers. If the genes lie in different chromosomes, as do the genes for whiteness and hairlessness, the two classes are equal.

Now suppose we have made a cross of this kind and find only 40 recombinations out of 100 mice. The deviation is twice its standard error, so we might regard it as significant. However, a mouse contains 19 pairs of chromosomes, (not counting the sex chromosomes), so there is an *a priori* probability of only about $\frac{1}{19}$ in favor of linkage, and the probability of actually finding it, after making the above observation, is only of the order of $\frac{2}{3}$. Let us see what we may expect if we repeat our experiment, not on 100 mice, but on 10,000, as Castle actually did in one particular case. In about one-third of such experiments we should find no linkage. The number of mice of new types not found in the ancestry will be 5,000 with a standard deviation of 50. In the other

cases there will be linkage. The number of mice of new types will be 4,000 with a standard deviation of 500. If we could breed a number of samples of a million mice the distributions of our samples would form something like a spectrum, with a bright line of small but finite breadth round 500,000, and a continuous distribution with a maximum at 400,000.

This example is interesting because it shows that in these biological problems *a priori* probability must be taken seriously. In fact I think that a concrete discussion of them may yet lead mathematicians to something better than Bayes' theorem. It is also interesting as showing that you can get a distribution like that of frequencies in a spectrum without quantum mechanics, in fact with very little but common sense. It is quite easy to devise an observation which yields several "bright lines," with or without a "continuous spectrum."

A second example is as follows. In man we can seldom observe more than two generations, but we desire to find linkages. Bernstein posed the following problem (I consider only the simplest of his ten cases). Suppose a man with brachydactyly and belonging to blood-group AB marries a woman with normal fingers belonging to group O. They have a number of children. Half receive the gene for brachydactyly, half for normal fingers. Also half receive the gene for the A agglutinogen, half that for the B. From a number of such samples we want to know if the genes are linked. But we do not know whether the genes for A and brachydactyly came into the father from the same parent or not. So we do not know whether, if linkage exists, we are to look out for an excess of A brachydactyls and B normals, or an excess of the other two types. In a series of such families we merely expect that there will be deviations on both sides of the mean value of one half.

Let me put the problem in terms more familiar to mathematicians. We have two bags containing black and white balls. One contains a proportion p of white, the other a proportion p of black. Samples are taken from these bags, but we do not know from which, or even whether it is equally likely that a sample has been taken from each bag. In fact we have a problem not in inverse probability but in double inverse probability. In this particular case the best value of p may be found by multiplying the numbers of black and white balls in each sample.

But a simple extension of Bernstein's problem demands the consideration of 9 types of bag, some of which may contain balls of as many as 12 colors. In such cases we must choose our function of the observed numbers so as to maximize the amount of information concerning the

linkage value, or what comes to the same thing, minimize its standard error. This is a fairly difficult problem, demanding both algebraic skill and appreciation of the biological situation.

Both the problems here stated are essentially problems of economy. How are we to get the most out of our data, or how are we to arrive at a given degree of precision with the least work? How many mice must we breed to be reasonably sure of our linkage? How many families must we observe? A mathematician is often impressed by the amount of time and energy which experimenters waste. The converse is however also true.

I have tried to show in this lecture, how biology and mathematics are stimulating one another. Their relation can only remain healthy so long as the stimulation is mutual. Theory should not run too far in front of experiment and observation, nor should it fail to tackle problems on which adequate data exist. These desiderata are best met when both are combined by such men as Henderson of Harvard, Lotka of the Metropolitan Life Assurance Company, and Wright of Chicago.

REFERENCES

FISHER, R. A. Statistical Methods for Research Workers.
FISHER, R. A. The Genetical Theory of Natural Selection.
HALDANE, J. B. S. The Causes of Evolution.
HENDERSON, L. J. Blood.
LOTKA, H. J. Physical Biology.

APPLICATIONS OF THE FOURIER TRANSFORM THEOREM

By Albert E. Heins

In this note it will be shown that the Fourier Transform Theorem may be used to unify the methods of solutions of linear partial differential equations of the second order with constant coefficients.

The Fourier Transform Theorem. Given $F(x_1 \cdots x_k)$ a function of k real variables which may take on complex values. Moreover let $F(x_1, \cdots x_k)$ be piecewise continuous, absolutely integrable, bounded, and have only a finite number of finite discontinuities. Under these hypotheses the Fourier Transform of exists and is given by

$$g(p_1, \cdots p_k) = \frac{1}{(2\pi)^{\frac{k}{2}}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} F(x_1 \cdots x_k) e^{-i \sum_{\lambda=1}^{k} p_\lambda x_\lambda} \, dx_1 \cdots dx_k$$

The transform of $g(p_1 \cdots p_k)$ is given by the dual relation

$$F(x_1 \cdots x_k) = \left(\frac{1}{2\pi}\right)^{\frac{k}{2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(p_1 \cdots p_k) e^{i \sum_{\lambda=1}^{k} p_\lambda x_\lambda} \, dp_1 \cdots dp_k$$

Henceforth we shall refer to the x_λ's, $\lambda = 1, \cdots k$ as coördinates in "coördinate space" and the p_λ's, $\lambda = 1, \cdots k$ as coördinates in "frequency space."

The results that are to follow will depend on a theorem closely related to the Fourier Transform Theorem. If $g_1(p_1 \cdots p_k)$ has a transform $f_1(x_1 \cdots x_k)$ and $g_2(p_1 \cdots p_k)$ has a transform $f_2(x_1 \cdots x_k)$ then the transform of $g_1(p_1 \cdots p_k) g_2(p_1 \cdots p_k)$ is given by

$$f(y_1 \cdots y_k) = \left(\frac{1}{2\pi}\right)^{\frac{k}{2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_1(x_1 \cdots x_k)$$

$$f_2(y_1 - x_1, \cdots y_k - x_k) \, dx_1 \cdots dx_k$$

This is called the Faltung Theorem.[1]

[1] Bochner, "Vorlesungen über Fouriersche Integrale," Leipzig, 1932.

The Wave Equation in Two Dimensions. Let us consider here the solution of

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}$$

By the Fourier Transform Theorem it is seen that

$$\frac{\partial u}{\partial x} \text{ corresponds to } ip_1 g$$

and similarly

$$\frac{\partial^2 u}{\partial x^2} \text{ corresponds to } -p_1^2 g$$

u and g depend on a parameter t (time) and hence it follows that

$$\frac{\partial^2 g}{\partial t^2} = -c^2(p_1^2 + p_2^2)g$$

g is then given by

$$g = A \cos c\sqrt{p_1^2 + p_2^2}\, t + B \sin c\sqrt{p_1^2 + p_2^2}\, t$$

For initial conditions we employ the following.

$$u(x, y, 0) = f(x, y): \qquad \frac{\partial u}{\partial t}(x, y, 0) = \phi(x, y)$$

Let the respective transforms of f and ϕ be $\Phi(p_1, p_2, 0)$ and $\frac{\partial \Phi}{\partial t}(p_1, p_2, 0)$. Then

$$g(p_1, p_2, t) = \Phi_{t=0} \cos c\sqrt{p_1^2 + p_2^2}\, t + \frac{1}{c\sqrt{p_1^2 + p_2^2}} \left.\frac{\partial \Phi}{\partial t}\right|_{t=0} \sin c\sqrt{p_1^2 + p_2^2}\, t$$

To evaluate $g(p_1, p_2, t)$ the transforms of $\cos c\sqrt{p_1^2 + p_2^2}\, t$ and

$$\frac{\sin c\sqrt{p_1^2 + p_2^2}\, t}{c\sqrt{p_1^2 + p_2^2}}$$

must be considered first. The functions $\cos c\sqrt{p_1^2 + p_2^2}\, t$ and

$$\frac{\sin c\sqrt{p_1^2 + p_2^2}\, t}{c\sqrt{p_1^2 + p_2^2}}$$

do not belong to the class of functions considered by the Fourier Transform Theorem. It is therefore necessary to consider the transforms of such functions as[2]

$$e^{-\epsilon\sqrt{p_1^2+p_2^2}}\cos c\sqrt{p_1^2+p_2^2}\,t$$

and

$$\frac{e^{-\epsilon\sqrt{p_1^2+p_2^2}}\sin c\sqrt{p_1^2+p_2^2}\,t}{c\sqrt{p_1^2+p_2^2}}.$$

which do belong to the class of functions considered and then to find the behaviour of the function as ϵ tends to zero.

Introduce now the following change of variables. Let the length of the vector from the origin to a point in coördinate space be γ and that of frequency space be ρ. Moreover let the angle between these two vectors be θ. Then the Fourier transform of $e^{-\epsilon\rho}\cos c\rho t$ becomes[3]

$$\int_0^{2\pi}\int_0^{\infty} e^{-\epsilon\rho}\cos c\rho t\, e^{i\rho\gamma\cos\theta}\,\rho\,d\theta\,d\rho$$

Upon integrating with respect to ρ we get

$$\frac{1}{2}\int_0^{2\pi} d\theta\left[\frac{\rho e^{-\epsilon\rho+i\rho(ct+\gamma\cos\theta)}}{-\epsilon+i(ct+\gamma\cos\theta)}-\frac{e^{-\epsilon\rho+i\rho(ct+\gamma\cos\theta)}}{[-\epsilon+i(ct+\gamma\cos\theta)]^2}\right.$$
$$\left.+\frac{\rho e^{-\epsilon\rho+i\rho(-ct+\gamma\cos\theta)}}{-\epsilon+i(-ct+\gamma\cos\theta)}-\frac{e^{-\epsilon\rho+i\rho(-ct+\gamma\cos\theta)}}{[-\epsilon+i(-ct+\gamma\cos\theta)]^2}\right]_0^{\infty}$$

or

$$\frac{1}{2}\int_0^{2\pi} d\theta\left[-\frac{1}{[-\epsilon+i(ct+\gamma\cos\theta)]^2}-\frac{1}{[-\epsilon+i(-ct+\gamma\cos\theta)]^2}\right]$$

As ϵ tends to zero this becomes

$$\frac{1}{2}\int_0^{2\pi} d\theta\left[\frac{1}{(ct+\gamma\cos\theta)^2}+\frac{1}{(ct-\gamma\cos\theta)^2}\right]=\frac{2\pi ct}{(c^2t^2-\gamma^2)^{3/2}}$$

Now consider the transform of

$$\frac{e^{-\epsilon\rho}\sin\rho ct}{c\rho}$$

[2] The method was suggested by Professor Wiener at a seminar in mathematics at the Massachusetts Institute of Technology.

[3] $e^{-\epsilon\rho}$ has been introduced as a convergence factor, where ϵ is an arbitrary positive quantity.

Then

$$\int_0^{2\pi}\int_0^{\infty} \frac{e^{-\epsilon\rho}\sin \rho ct\, e^{i\gamma\rho\cos\theta}}{\rho c}\,\rho\, d\theta\, d\rho$$

$$\frac{1}{2ic}\int_0^{2\pi} d\theta\left[\frac{1}{-\epsilon + i(ct+\gamma\cos\theta)} - \frac{1}{-\epsilon - i(ct-\gamma\cos\theta)}\right]$$

Letting ϵ tend to zero and integrating over θ one gets

$$-\frac{2\pi}{c\sqrt{c^2t^2-\gamma^2}}$$

Applying the Faltung theorem we get

$$u(x,y,t) = \frac{2\pi c}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{t f(x',y')\,dx'\,dy'}{[\sqrt{c^2t^2-(x-x')^2-(y-y')^2}]^3} - \frac{2\pi c}{(2\pi)^2 c}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\phi(x',y')\,dx'\,dy'}{\sqrt{c^2t^2-(x-x')^2-(y-y')^2}}$$

$$= -\frac{1}{2\pi c}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\partial}{\partial t}\frac{f(x',y')\,dx'\,dy'}{\sqrt{c^2t^2-(x-x')^2-(y-y')^2}} - \frac{1}{2\pi c}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\phi(x',y')\,dx'\,dy'}{\sqrt{c^2t^2-(x-x')^2-(y-y')^2}}$$

This is the solution that Goursat[4] obtains by much less direct methods.[5]

It is interesting to remark here that if the integration has first been carried out with respect to θ then the problem would reduce to the evaluation of integrals of the form

$$3\pi\int_0^{\infty}\rho J_0(\gamma\rho)\,e^{-\epsilon\rho}\cos c\rho t\,d\rho: \qquad 2\pi\int_0^{\infty}J_0(\gamma\rho)\,e^{-\epsilon\rho}\sin c\rho t\,d\rho$$

where the $J's$ are Bessels functions. This would make the calculations a little more difficult to carry through.[6]

Now the calculations which have been indicated not only apply to the

[4] Goursat, "Cours D'Analyse," Tome III, page 103, Paris, 1927.

[5] By methods somewhat analogous to these Webster derives the same results. "Partial Differential Equations of Mathematical Physics," G. E. Stechert, New York, 1933.

[6] It is interesting to note that if one integrates with respect to θ in a space of an even number of dimensions one gets Bessels functions, while in a space of an odd number of dimensions one gets trigonometric functions.

two dimensional wave equations, but to the n dimensional wave equation. That is we can consider by these methods the solution of

$$\sum_{\lambda=1}^{k} \frac{\partial^2 u}{\partial x_\lambda^2} = \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2}$$

The author wishes to take this occasion to point out that the Fourier Transform may be also used to derive the inhomogeneous solution of partial differential equations with constant coefficients. Let us consider here the Poisson equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = -4\pi\omega(x, y, z)$$

By the Fourier Transform Theorem we have

$$(p_1^2 + p_2^2 + p_3^2)g = 4\pi\Omega(p_1, p_2, p_3)$$

where $g(p_1, p_2, p_3)$ is the transform of $u(x, y, z)$ and $\Omega(p_1, p_2, p_3)$ is the transform of $\omega(x, y, z)$. Then

$$u(x, y, z) = \left(\frac{1}{2\pi}\right)^{3/2} \int\int_{-\infty}^{\infty}\int \frac{4\pi\Omega e^{i(p_1 x + p_2 y + p_3 z)}}{p_1^2 + p_2^2 + p_3^2} dp_1 dp_2 dp_3$$

Introducing polar coördinates for the p's, we have

$$u(x, y, z) = \left(\frac{1}{2\pi}\right)^{3/2} \int_0^{2\pi}\int_0^{\infty}\int_0^{\pi} \frac{4\pi\Omega e^{i\gamma\rho\cos\theta}\rho^2 \sin\theta d\phi d\rho d\theta}{\rho^2}$$

Consider the transform of $\frac{1}{\rho^2}$. We have then

$$\left(\frac{1}{2\pi}\right)^{3/2} \int_0^{2\pi}\int_0^{\infty}\int_0^{\pi} e^{i\gamma\rho\cos\theta} \sin\theta d\phi d\rho d\theta$$

$$= \left(\frac{1}{2\pi}\right)^{3/2} \int_0^{2\pi}\int_0^{\infty} \frac{2\sin\gamma\rho\, d\phi\, d\rho}{\gamma\rho} = \frac{(2\pi)^{1/2}}{\gamma}$$

Then by the Faltung theorem, we have

$$u(x, y, z) = \frac{1}{4\pi} \int\int_{-\infty}^{\infty}\int \frac{\omega(x', y', z')\, dx'\, dy'\, dz'}{\sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}}$$

This is a more direct method of getting the solution to this equation, than the one usually employed.

If we consider linear partial differential equations with variable

coefficients, we may pass into an integral equation, provided the coefficients satisfy the conditions of the transform theorem. Consider for example Schrodinger's equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + (V - E)u = 0$$

where V is the potential function, and u is the wave function. Then by the Transform Theorem we have

$$g(p_1, p_2, p_3) = \frac{1}{\rho^2 + E} \int\int_{-\infty}^{\infty} \int v(p_1 - p_1', p_2 - p_2', p_3 - p_3') \, g(p_1', p_2', p_3') \, dp_1' \, dp_2' \, dp_3'$$

where $g(p_1, p_2, p_3)$ is the transform of $u(x, y, z)$ and $v(p_1, p_2, p_3)$ is the transform of V, the potential function. By means of the theory of integral equations we can at least find the characteristic values of E or what they correspond to physically, the correct energy levels. This will be useful when one cannot solve Schrodinger's equations directly.

ON S, P, D AND F BIHARMONIC POLYNOMIALS

By Frank L. Hitchcock

1. **Biharmonic polynomials.** Consider polynomials homogeneous of degree l in x, y, z and also homogeneous in another set of variables x_1, y_1, z_1, but of degree l' which may be either the same as l or different. Such a polynomial $\psi(x, y, z, x_1, y_1, z_1)$ will be called *biharmonic* if it satisfies Laplace's equation separately and independently in each set of variables; that is if

$$\Delta\psi = 0 \text{ and } \Delta_1\psi = 0 \tag{1}$$

where

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad \text{and} \quad \Delta_1 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial y_1^2} + \frac{\partial^2}{\partial z_1^2}. \tag{2}$$

The following further notation will be useful:

$$r^2 = x^2 + y^2 + z^2 \text{ and } r_1^2 = x_1^2 + y_1^2 + z_1^2; \tag{3}$$

H_n will denote a polynomial homogeneous of degree n in x, y, z and satisfying Laplace's equation for these variables but free of the other variables x_1, y_1, z_1. Similarly K_{n_1} will denote a polynomial homogeneous of degree n_1 in x_1, y_1, z_1 satisfying Laplace's equation in x_1, y_1, z_1 but not containing x, y, z. In other words H and K denote ordinary spherical harmonics in their respective variables.

The following theorem follows readily from known properties of spherical harmonics:

Theorem 1. Any polynomial $P(x, y, z; x_1, y_1, z_1)$ homogeneous of degree l in x, y, z and homogeneous of degree l' in x_1, y_1, z_1 can be expanded in the form

$$P = \psi_{l,l'} + r^2\psi_{l-2,l'} + r^2r_1^2\,\psi_{l-2,l'-2} + \cdots + r^{2s}r_1^{2s'}\psi_{l-2s,l'-2s'} + \cdots \tag{4}$$

where the polynomials ψ are biharmonic of degrees indicated by their subscripts.

For consider any term of P, namely $x^iy^jz^kx_1^uy_1^vz_1^w$. The factor $x^iy^jz^k$ can be expanded as

$$x^i\,y^j\,z^k = H_l + r^2H_{l-2} + \cdots + r^{2s}H_{l-2s} + \cdots \tag{5}$$

by a well-known theorem.[1] Here the successive polynomials H are ordinary harmonics of degrees indicated by their subscripts. Similarly $x_1^u y_1^v z_1^w$ can be expanded in the form

$$x_1^u y_1^v z_1^w = K_{l'} + r_1^2 K_{l'-2} + \cdots + r_1^{2s} K_{l'-2s} + \cdots \tag{6}$$

where the successive polynomials K are ordinary spherical harmonics in the variables x_1, y_1, z_1 and are of the degrees indicated by their subscripts. Expanding all terms of P in this way and collecting the coefficients of the powers of r and of r_1, each such coefficient is a sum of products HK and therefore biharmonic.

When occasion arises actually to make the expansion (4) it is most natural to use the method of Dougall (Hobson, p. 148). To apply this method to a polynomial in two sets of variables we begin by taking note of the fact that, when P_n is any polynomial homogeneous in x, y, z and of degree n in these variables (whether it is also a function of other variables or not), we have identically.

$$\Delta(r^k P_n) = C_n^k r^{k-2} P_n + r^k \Delta P_n; \tag{7}$$

the constant C_n^k is given by

$$C_n^k = k(2n + k + 1). \tag{8}$$

It follows that, in general, starting with any P_n, we can obtain a harmonic polynomial $H[P_n]$ by the following formula

$$H[P_n] = P_n - \frac{r^2 \Delta P_n}{C_{n-2}^2} + \frac{r^4 \Delta^2 P_n}{C_{n-2}^2 C_{n-4}^4} - \cdots + (-1)^j \frac{r^{2j} \Delta^j P_n}{C_{n-2}^2 C_{n-4}^4 \cdots C_{n-2j}^{2j}} + \cdots \tag{9}$$

where

$$\Delta^j = \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right]^j; \tag{10}$$

exception occurs when and only when P_n contains r as a factor, in which case the right member of (9) vanishes identically. We now apply Dougall's method to $P(x, y, z; x_1, y_1, z_1)$, assuming degrees l and l' as in

[1] Hobson, *Spherical and Ellipsoidal Hormonics*. p. 147.

Theorem 1; namely we write the identity

$$P \equiv H[P] + \frac{r^2 H[\Delta P]}{C^2_{l-2}} + \frac{r^4 H[\Delta^2 P]}{C^2_{l-4} C^4_{l-4}} + \cdots + \frac{r^{2s} H[\Delta^s P]}{C^2_{l-2s} C^4_{l-2s} \cdots C^{2s}_{l-2s}} + \cdots . \tag{11}$$

The coefficients of the successive powers of r^2 on the right will be harmonic in x, y, z but not in general in x_1, y_1, z_1.

Now however we apply the Dougall expansion to $H[P]$, and also to $H[\Delta P]$, $H[\Delta^2 P]$, etc., using Δ_1 instead of Δ and r_1 instead of r; also K instead of H (see (13) below). In carrying out the work it is well to use the abbreviated notation

$$D(l,j) = C^2_{l-2}\, C^4_{l-4} \cdots C^{2j}_{l-2j} = 2.4.6 \cdots 2j.\,(2l-1)(2l-3) \cdots (2l-2j+1). \tag{12}$$

We then note that, in strict analogy with (9), if P_{n_1} is any polynomial homogeneous of degree n_1 in x_1, y_1, z_1, whether a function of other variables or not, we can obtain from it a polynomial harmonic in x_1, y_1, z_1 by the formula

$$K\{Pn_1\} = Pn_1 - \frac{r_1^2 \Delta_1 Pn_1}{D(n_1,1)} + \frac{r_1^4 \Delta_1^2 Pn_1}{D(n_1,2)} - \cdots + (-1)^{j'} \frac{r_1^{2j'} \Delta_1^{j'} Pn_1}{D(n_1,j')} + \cdots \tag{13}$$

with exception when and only when P_{n_1} contains r_1 as a factor. We now substitute $H[P]$ for P_{n_1} in every term; the result is biharmonic, for $H[P]$ is harmonic in x, y, z, hence $\Delta_1 H[P], \cdots, \Delta_1^{j'} H[P]$ are all harmonic in x, y, z; and the expression as a whole is harmonic in x_1, y_1, z_1. Calling this biharmonic $\psi(P)$, we carry out the substitution by aid of (9), with $n = l$ and $n_1 = l'$. The result is

$$\psi(P) = K\{H[P]\} = P - \frac{r^2 \Delta P}{D(l,1)} - \frac{r_1^2 \Delta_1 P}{D(l',1)} + \cdots + (-1)^{j+j'} \frac{r^{2j} r_1^{2j'} \Delta^j \Delta_1^{j'} P}{D(l,j)\, D(l',j')} + \cdots \tag{14}$$

in which we are to take $D(l,0) = D(l',0) = 1$, and the summation is for both j and $j' = 0, 1, 2, \cdots$. The expansion will terminate when j

and j' exceed $l/2$ and $l'/2$ respectively. Since Δ and Δ_1 are commutative operations the right side is formally symmetrical in the operators H and K, in the sense that $K\{H[P]\} = H[K\{P\}] = \psi(P)$. Formula (14) shows how, from any polynomial homogeneous in two sets of variables x, y, z and x_1, y_1, z_1, provided neither r^2 nor r_1^2 is a factor, we may obtain a biharmonic $\psi(P)$.

We are now ready to develop $H[P]$ by Dougall's method as was purposed. Much space will be saved by adopting the further notation

$$\begin{aligned} E(l, s) &= C^2_{l-2s}C^4_{l-2s} \cdots C^{2s}_{l-2s} \\ &= 2.4. \cdots 2s.\ (2l - 4s + 3)(2l - 4s + 5) \cdots (2l - 2s + 1), \end{aligned} \tag{15}$$

with $E(l, 0) = 1$. With variables x_1, y_1, z_1 the Dougall development for $H[P]$ is then

$$\begin{aligned} H[P] \equiv K\{H[P]\} &+ \frac{r_1^2 K\{\Delta_1 H[P]\}}{E(l', 1)} \\ &+ \frac{r_1^4 K\{\Delta_1^2 H[P]\}}{E(l', 2)} + \cdots + \frac{r_1^{2s'} K\{\Delta_1^{s'} H[P]\}}{E(l', s')} + \cdots . \end{aligned} \tag{16}$$

Inspection of (9) shows that Δ_1 and H are commutative, so that

$$K\{\Delta_1^{s'} H[P]\} = K\{H[\Delta_1^{s'} P]\} = \psi(\Delta_1^{s'} P) \tag{17}$$

and (16) may be written more simply as

$$\begin{aligned} H[P] \equiv \psi(P) &+ \frac{r_1^2 \psi(\Delta_1 P)}{E(l', 1)} \\ &+ \frac{r_1^4 \psi(\Delta_1^2 P)}{E(l', 2)} + \cdots + \frac{r_1^{2s'} \psi(\Delta_1^{s'} P)}{E(l', s')} + \cdots . \end{aligned} \tag{18}$$

The desired expansions for $H[\Delta P]$, $H[\Delta^2 P]$, etc. are similar, merely having ΔP, $\Delta^2 P$, $\cdots$, respectively, in place of P. Substituting these expansions in (11) gives

$$\begin{aligned} P \equiv \psi(P) &+ \frac{r^2 \psi(\Delta P)}{E(l, 1)} \\ &+ \frac{r_1^2 \psi(\Delta_1 P)}{E(l', 1)} + \cdots + \frac{r^{2s} r_1^{2s'} \psi(\Delta^s \Delta_1^{s'} P)}{E(l, s)\, E(l', s')} + \cdots \end{aligned} \tag{19}$$

which is the desired identity of form (4) and shows how the terms of (4) may be found by direct operations: the linear operator ψ is defined by (14).

The manifold of all polynomials homogeneous of degree l in x, y, z is a unitary space provided orthogonality is defined with reference to integration over the surface of a unit sphere with centre at the origin. Then $H[P_l]$ as defined by (9) is the harmonic component of P_l, and similarly for $K\{P_{l'}\}$. In the same sense $\psi(P)$ is the biharmonic component of P in the manifold of all polynomials homogeneous in x, y, z and in x_1, y_1, z_1 of respective degrees l and l', orthogonality being now defined with reference to integration over the unit sphere in both sets of variables. The Dougall expansion may be thought of as an identical operation. That expansion of a given P in form (4) is unique follows from the orthogonality of the terms of (4) to one another: two harmonics of different degrees are always orthogonal.

The identity (19) may also be viewed in a more elementary way: it subsists by virtue of various identical relations among the constants C_n^k which may be obtained by expanding each term by (14) and equating to zero the coefficients of the powers of Δ and Δ_1.

The generalization of (14) and (19) to three or more sets of three variables each is obvious. They may also be generalized to harmonics in sets of more than three variables, as well as to linear operators other than Δ whose characteristic roots are known.

2. **Primitive submanifolds.** Consider the manifold of harmonic polynomials H_l homogeneous of degree l in x, y, z but not containing x_1, y_1, z_1. There are several well-known ways to choose a basis of linearly independent functions in the manifold. For the purpose of this paper the following will be the most convenient: let

$$H_l^m = (x+iy)^m Z_l^m, \quad (m = l, l-1, \cdots, +1, 0, -1, \cdots, -l+1, -l), \tag{20}$$

where

$$\begin{aligned} Z_l^m = z^{l-m} &- \frac{(l-m)(l-m-1)}{2(2l-1)} r^2 z^{l-m-2} \\ &+ \cdots + (-1)^j \frac{(l-m)\cdots(l-m-2j+1)}{D(l,j)} r^{2j} z^{l-m-2j} + \cdots . \end{aligned} \tag{21}$$

That the $(2l+1)$ functions thus defined are harmonic is well-known but can be seen directly from the fact that they are the expansions of $(x+iy)^m z^{l-m}$ by (9). Negative and positive values of m are related by the identity

$$(x+iy)^{-|m|} Z_l^{-|m|} = (-1)^{|m|} (x-iy)^{+|m|} Z_l^{+|m|}. \tag{22}$$

The H_l^m are mutually orthogonal but not normalized. (Normalization factors will be given below). Basis functions $K_{l'}^{m'}$ of x_1, y_1, z_1 are defined in a manner precisely similar.

Since the H_l^m and the $K_{l'}^{m'}$ are linearly independent so also are the $(2l + 1)(2l' + 1)$ products $H_l^m K_{l'}^{m'}$ and this is the maximum number of linearly independent biharmonics $\psi_{l,l'}$. For any given $\psi_{l,l'}$ can be written as a linear combination of the $H_l^m K_{l'}^{m'}$: assume ψ to be biharmonic and write

$$\psi = \sum_{i\,j\,k} x^i y^j z^k P_{ijk}(x_1, y_1, z_1) \tag{23}$$

where the P_{ijk} are polynomials in x_1, y_1, z_1 only, obtained by collecting all terms in which i, j, and k are constant. The P_{ijk} must be harmonic in x_1, y_1, z_1 because

$$\Delta_1 \psi = \sum_{i\,j\,k} x^i y^j z^k \Delta_1 P_{ijk} = 0 \tag{24}$$

which is impossible unless *each* $\Delta_1 P_{ijk}$ vanishes; for the $x^i y^j z^k$ are linearly independent. The P_{ijk} may therefore be written as linear combinations of the $K_{l'}^{m'}$ and we may collect together terms in which m' is constant so as to write

$$\psi = \sum_{m'} P_{m'}(x, y, z)\, K_{l'}^{m'} \tag{25}$$

where the $P_{m'}$ must be harmonic in x, y, z since the $K_{l'}^{m'}$ are linearly independent, and may be written as linear combinations of the H_l^m. We thus obtain ψ as a linear combination of the products $H_l^m K_{l'}^{m'}$.

Subjecting both sets of variables to the *same* orthogonal linear transformation with determinant $+1$ will be called a *rotation*. The transform of ψ will be written $\mathbf{R}\psi$. If the transform of every function of a manifold belongs to the manifold we say the manifold is *invariant*. A manifold is *primitive* (or *irreducible*) if it contains no proper invariant submanifold as the transformation in question takes on all possible values. It is well-known that the manifold H_l of ordinary harmonics of degree l is primitive under rotation. The manifold of biharmonics $\psi_{l,l'}$ of given degrees l, l' is invariant but not primitive. It is known that *any* manifold which is primitive under rotation is isomorphic with one of the manifolds of ordinary harmonics: this means that the number of linearly independent functions in a primitive manifold is odd, say $2L + 1$, and can be so chosen as to transform under rotation like the functions (20) when, in (20), l is put equal to L. Furthermore the fact that, for given l

and l', and with l' not greater than l, we have always a set of $2l' + 1$ primitive submanifolds in the manifold of the $\psi_{l,l'}$ with

$$L = l + l', \quad l + l' - 1, \quad l + l' - 2, \cdots, \quad l - l', \tag{26}$$

is also well-known.[2] Borrowing the convenient terminology of Physics we may say that a primitive manifold with $L = 0$ is an S manifold, and similarly those with $L = 1, 2, 3$, etc. are P, D and F manifolds, respectively. In the case of harmonics in x, y, z alone there is just one primitive manifold for each value of l, namely the linear combinations of the H_l^m. With biharmonics the situation is quite different: inspection of (26) shows that there exists an infinite series of S manifolds one for each value of l, since we may always take $l' = l$ and $L = l - l' = 0$. And there are two distinct and quite unlike infinite series of P manifolds: first, by taking $l' = l$ and $L = l - l' + 1 = +1$, there is a P manifold corresponding to each value of l except $l = 0$; and second, by taking $l' = l - 1$ and $L = l - l' = +1$, there is a P manifold corresponding to each value of l if l exceeds $+1$. Similarly there are three infinite series of D manifolds, four infinite series of F manifolds, and so on. It will appear below that these various series of primitive manifolds are quite unlike in character, as may well be anticipated from the physical analogy, even though, in the present state of our knowledge, this analogy is only qualitative or perhaps at times semi-quantitative.

What is most important for the study of biharmonics is that each primitive manifold is unique in the sense that with given l, l', and L the basis functions which shall transform under rotation like a set of $2L + 1$ functions of the form (20) are unique except for a constant factor. Such would not be the case with polynomials harmonic in three or more sets of variables, where, in analogous circumstances, we may have two or more primitive manifolds having the same value of L and of the same degrees in each set of variables, and the decomposition into primitive manifolds may be made in more than one way.

3. **Infinitesimal rotations.** With given l, l', and L, a set of functions which transforms under rotation like a set of functions (20) where, in (20), we set L in place of l, will be denoted by $\psi_L^M(l, l')$, wherein $M = L$, $L - 1, \cdots, -L$. The explicit formulation of these functions is greatly facilitated by the use of the differential operators commonly called *infinitesimal rotations*; strictly speaking this designation is a misnomer, for these operators are not special cases of the general rota-

[2] H. Weyl, *Gruppentheorie und Quantenmechanik, Erste Auflage*, p. 133, 134.

tion **R** but in fact yield, to the first order of small quantities, the *change* induced in a function by a rotation through a small angle. If $\mathbf{R}_z$ denotes rotation about the axis of Z through an angle φ an easy calculation shows that, P being a polynomial in x, y, z,

$$\mathbf{R}_z P = P + \varphi \left[y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right] P + \text{terms in } \varphi^2, \varphi^3, \text{ etc.} \tag{27}$$

and similarly for other axes. We write first

$$D_z = y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y}, \qquad D_x = z \frac{\partial}{\partial y} - y \frac{\partial}{\partial z}, \qquad D_y = x \frac{\partial}{\partial z} - z \frac{\partial}{\partial x} \tag{28}$$

then define two new operators

$$D_\xi = -i [D_x + iD_y], \qquad D_\eta = -i [D_x - iD_y]. \tag{29}$$

We then have

$$\begin{gathered} D_\xi H_l^m = (l - m) H_l^{m+1}; \qquad D_\eta H_l^m = (l + m) H_l^{m-1}; \\ D_z H_l^m = -mi H_l^m; \end{gathered} \tag{30}$$

as is easily seen by taking as independent variables r, z, and ξ $(= x + iy)$, when D_ξ becomes $\xi \frac{\partial}{\partial z}$ and D_z becomes $-i\xi \frac{\partial}{\partial \xi}$; then by taking as independent variables r, z, and $\eta(= x - iy)$ we find $D_\eta = -\eta \frac{\partial}{\partial z}$. The simplicity of the relations (30) is one of the reasons why in the present investigation the polynomials (20) are taken as basis.

When P is a function of x_1, y_1, z_1 as well as of x, y, z and the same rotation affects both sets of variables simultaneously, an expansion like (27) shows that the coefficient of φ now is $D_z + D_{z_1}$ into P, where $D_{z_1} = y_1 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial y_1}$; and similarly for the other infinitesimal operators. Since by hypothesis the ψ_L^M transform under rotation like the H_l^m we shall have, analogous with (30),

$$\begin{gathered} (D_\xi + D_{\xi_1}) \psi_L^M = (L - M) \psi_L^{M+1}; \quad (D_\eta + D_{\eta_1}) \psi_L^M = (L + M) \psi_L^{M-1}; \\ (D_z + D_{z_1}) \psi_L^M = -iM \psi_L^M. \end{gathered} \tag{31}$$

The relation of ψ_L^M to each other and to infinitesimal rotation is clearly seen if these functions be indicated as elements of a matrix whose rows have a constant value of L and whose columns have a constant value of M. When $l' < l$ the matrix is trapezoidal,

$$
\begin{array}{ccccccccc}
\psi^{i+i',}_{i+i',} & \psi^{i+i',}_{i+i'-1}, & \psi^{i+i',}_{i+i'-2}, & \cdots, & \psi^{i+i',}_{i-i'+1}, & \psi^{i+i',}_{i-i'}, & \cdots, & \psi^{i+i',}_{-i-i'+1}, & \psi^{i+i',}_{-i-i'} \\
\psi^{i+i'-1,}_{i+i'-1}, & \psi^{i+i'-1,}_{i+i'-2}, & \cdots, & \psi^{i+i'-1,}_{i-i'+1}, & \psi^{i+i'-1,}_{i-i'}, & \cdots, & \psi^{i+i'-1,}_{-i-i'+2}, & \psi^{i+i'-1,}_{-i-i'+1}, & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & & \\
\psi^{i-i',}_{i-i'}, & \psi^{i-i',}_{i-i'-1}, & \cdots, & \psi^{i-i',}_{-i+i'+1}, & \psi^{i-i',}_{-i+i'}, & \cdots, & \psi^{i-i',}_{-i-i'+1}, & \psi^{i-i',}_{-i-i'} &
\end{array}
\tag{32}
$$

but when $l - l' = 0$ the lower row evidently contains only the single element ψ_0^0. It is to be understood that l and l' are given and constant throughout. The elements of any one row are a basis for a primitive submanifold. The effect of the operator $D_\eta + D_{\eta_1}$ is to shift any element ψ_L^M one place to the right and multiply it by $(L + M)$. Similarly $D_\xi + D_{\xi_1}$ shifts any element one place to the left and multiplies it by $(L - M)$. The elements which are thus pushed off the diagram are annulled. $D_z + D_{z_1}$ leaves each element in position, but multiplies it by $-iM$, each element being an axis (eigenfunction) of this operator, with root $-iM$.

4. **Explicit formulation of primitive submanifolds.** The problem of calculating any desired polynomial $\psi_L^M(l, l')$ can now be readily solved. The following lemma is helpful,—

Lemma. Any $\psi_L^M(l, l')$, where L, M, l and l' are given and fixed, can be written as a linear combination of those products $H_l^m K_{l'}^{m'}$ for which $m + m' = M$.

Proof. In the manifold of biharmonic polynomials $\psi_{l,l'}$ the polynomials which satisfy the equation

$$(D_z + D_{z_1})\psi = -iM\psi \tag{33}$$

form a submanifold. The $\psi_L^M(l, l')$ having l, l' and M constant but L variable are by (31) a basis for this submanifold, i.e. the elements of that column of (32) in which the upper index is equal to M. But those products $H_l^m K_{l'}^{m'}$ in which $m + m' = M$ are also a basis for the same submanifold, because

$$(D_z + D_{z_1})H_l^m K_{l'}^{m'} = -i(m + m')H_l^m K_{l'}^{m'} \tag{34}$$

by (30). Hence the lemma.[3]

We begin with those ψ_L^M for which $L = M$, i.e. those which stand at the left extremity of their rows in (32). Evidently a ψ_L^L is uniquely determined, aside from a constant factor, by the two conditions

$$(D_z + D_{z_1})\psi_L^L = -iL\,\psi_L^L \text{ and } (D_\xi + D_{\xi_1})\psi_L^L = 0 \tag{35}$$

[3] A more elementary proof: the submanifold includes all functions which, with independent variables r, z, and $x + iy$, contain $(x + iy)^m(x_1 + iy_1)^{m'}$ with $m + m' = M$. Then consider the rotation $\mathbf{R}_z$. Still better, the lemma may be regarded as a very special case of the relations given by Slater, Theory of Complex Spectra, Phys. Rev. **34**, 1929, 1301, neglecting spin and symmetry and taking $H = 1$.

wherein L, as well as l and l', has a given value. For the first condition requires that ψ, as an element of (32), lies in the column with upper index equal to L, and the second condition requires ψ to be at the left of a row. The first condition will be automatically satisfied if we write

$$\psi_L^L(l, l') = H_l^{l-k}K_{l'}^{l'} + A_1 H_l^{l-k+1}K_{l'}^{l'-1} + \cdots + A_s H_l^{l-k+s}K_{l'}^{l'-s} + \cdots \tag{36}$$

where the A_s are undetermined coefficients and where

$$k = l + l' - L \tag{37}$$

so that, for successive rows of (32), k runs from 0 to $2l'$. In each term on the right of (36) we have $m + m' = L$. We now operate with $D_\xi + D_{\xi_1}$ using (30) and (31) and have

$$0 = kH_l^{l-k+1}K_{l'}^{l'} + A_1 H_l^{l-k+1}K_{l'}^{l'} + \cdots + [(k-s)A_s + (s+1)A_{s+1}]H_l^{l-k+s+1}K_{l'}^{l'-s} + \cdots \tag{38}$$

whence the A_s are binomial coefficients with alternating signs and we have, for the general $\psi_L^L(l, l')$

$$\psi_L^L = H_l^{l-k}K_{l'}^{l'} - kH_l^{l-k+1}K_{l'}^{l'-1} + \cdots + (-1)^s\binom{k}{s}H_l^{l-k+s}K_{l'}^{l'-s} + \cdots . \tag{39}$$

By (30) and (31) we may now obtain those elements of (32) in which M is less than L by successive operation with $D_\eta + D_{\eta_1}$. Any desired $\psi_L^M(l, l')$ may thus be calculated.

By this method, supposing l and l' given and fixed, we express each of the $(2l + 1)(2l' + 1)$ polynomials ψ_L^M as a linear combination of the $(2l + 1)(2l' + 1)$ polynomials $H_l^m K_{l'}^{m'}$ by means of $(2l + 1)(2l' + 1)$ linear equations. The coefficients will form a square matrix of $(2l + 1)(2l' + 1)$ rows and columns. By virtue of the lemma, however, this large matrix can be so partitioned as to consist of $2l + 2l' + 1$ small square matrices each with its main diagonal lying along the main diagonal of the large matrix: each small matrix comprises the coefficients from the equations which express the elements of one column of (32) in terms of those $H_l^m K_{l'}^{m'}$ for which $m + m'$ has the appropriate constant value, viz. the upper index in the column in question from (32).

As the simplest example, let $l = l' = 1$. Then

$$(2l + 1)(2l' + 1) = 9,$$

and $2l + 2l' + 1 = 5$. We shall expect a matrix of nine rows and columns partitioned diagonally into five small matrices; and three primitive submanifolds respectively S, P and D. The matrix (32) becomes for this case

$$\begin{matrix} \psi_2^2, \psi_2^1, \psi_2^0, \psi_2^{-1}, \psi_2^{-2} \\ \psi_1^1, \psi_1^0, \psi_1^{-1} \\ \psi_0^0 \end{matrix} \tag{40}$$

We may begin by putting $L = M = 0$ in (39) whence $k = 2$ by (37). Then (39) becomes

$$\psi_0^0(1, 1) = H_1^{-1}K_1^1 - 2H_1^0K_1^0 + H_1^1K_1^{-1}. \tag{41}$$

Next we may put $L = M = 1$ in (39) and $k = 1$ by (37). From (39) we now have

$$\psi_1^1(1, 1) = H_1^0K_1^1 - H_1^1K_1^0 \tag{42}$$

and by operating twice in succession with $D_{\eta} + D_{\eta_1}$, using (30) and (31),

$$\begin{aligned} 2\psi_1^0(1, 1) &= H_1^{-1}K_1^1 + 2H_1^0K_1^0 - 2H_1^0K_1^0 - H_1^1K_1^{-1} \\ &= H_1^{-1}K_1^1 - H_1^1K_1^{-1} \end{aligned} \tag{43}$$

$$2\psi_1^{-1}(1, 1) = 2H_1^{-1}K_1^0 - 2H_1^0K_1^{-1}. \tag{44}$$

The last three results determine a basis for the P manifold, i.e. the second row of (40). The D manifold will have five basis functions, viz. the upper row of (40). We put $L = M = 2$ in (39) with $k = 0$ by (37) and have from (39)

$$\psi_2^2(1, 1) = H_1^1K_1^1 \tag{45}$$

and operating four times in succession with $D_{\eta} + D_{\eta_1}$ using (30) and (31),

$$4\psi_2^1(1, 1) = 2H_1^0K_1^1 + 2H_1^1K_1^0 \tag{46}$$

$$\begin{aligned} 12\psi_2^0(1, 1) &= 2H_1^{-1}K_1^1 + 4H_1^0K_1^0 + 4H_1^0K_1^0 + 2H_1^1K_1^{-1} \\ &= 2H_1^{-1}K_1^1 + 8H_1^0K_1^0 + 2H_1^1K_1^{-1} \end{aligned} \tag{47}$$

$$\begin{aligned} 24\psi_2^{-1}(1, 1) &= 4H_1^{-1}K_1^0 + 8H_1^{-1}K_1^0 + 8H_1^0K_1^{-1} + 4H_1^0K_1^{-1} \\ &= 12H_1^{-1}K_1^0 + 12H_1^0K_1^{-1} \end{aligned} \tag{48}$$

$$\begin{aligned} 24\psi_2^{-2}(1, 1) &= 12H_1^{-1}K_1^{-1} + 12H_1^{-1}K_1^{-1} \\ &= 24H_1^{-1}K_1^{-1} \end{aligned} \tag{49}$$

Equations (41)–(49) give the nine elements of (40) in terms of the $H_1^m K_1^{m'}$. Arranging first in order of columns and, within each column, in descending order of values of L, the matrix of the coefficients is

	$H_1^1K_1^1$	$H_1^1K_1^0$	$H_1^0K_1^1$	$H_1^1K_1^{-1}$	$H_1^0K_1^0$	$H_1^{-1}K_1^1$	$H_1^0K_1^{-1}$	$H_1^{-1}K_1^0$	$H_1^{-1}K_1^{-1}$
ψ_2^2	1	·	·	·	·	·	·	·	·
ψ_2^1	·	$\frac{1}{2}$	$\frac{1}{2}$	·	·	·	·	·	·
ψ_1^1	·	-1	1	·	·	·	·	·	·
ψ_2^0	·	·	·	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$	·	·	·
ψ_1^0	·	·	·	$-\frac{1}{2}$	0	$\frac{1}{2}$	·	·	·
ψ_0^0	·	·	·	1	-2	1	·	·	·
ψ_2^{-1}	·	·	·	·	·	·	$\frac{1}{2}$	$\frac{1}{2}$	·
ψ_1^{-1}	·	·	·	·	·	·	-1	1	·
ψ_2^{-2}	·	·	·	·	·	·	·	·	1

(50)

The elements indicated by dots are zero.

5. **Normalization.** Normalizing factors for ordinary spherical harmonics are well known. With the H_l^m defined by (20) and (21) and writing $N_l^m = (H_l^m, H_l^m)$ for the integral of the square of the absolute value of H_l^m over the surface of the unit sphere it is easy to show, either from familar formulas or by direct computation that

$$N_l^m = (H_l^m, H_l^m) = \frac{4^{l+1}(l!)^2\pi}{(2l+1)!} \cdot \frac{1}{\binom{2l}{l-m}} \tag{51}$$

where the denominator of the second factor is a binomial coefficient. In this formula m is to be taken positive. Always $N_l^m = N_l^{-m}$. Writing

$$H_l^m = (N_l^m)^{\frac{1}{2}}\mathbf{H}_l^m \tag{52}$$

and similarly for $K_l^{m'}$ it is easy to write equations (50) with normalized **H** and **K**. For example we find from (51)

$$N_1^0 = \frac{4\pi}{3}, \qquad N_1^1 = \frac{8\pi}{3}, \tag{53}$$

so that from (50)

$$\psi_2^1 = \frac{1}{2}\left(\frac{8\pi}{3}\right)^{\frac{1}{2}}\left(\frac{4\pi}{3}\right)^{\frac{1}{2}}\mathbf{H}_1^1\mathbf{K}_1^0 + \frac{1}{2}\left(\frac{8\pi}{3}\right)^{\frac{1}{2}}\left(\frac{4\pi}{3}\right)^{\frac{1}{2}}\mathbf{H}_1^0\mathbf{K}_1^1 \tag{54}$$

and similarly in all cases. Finally let $\mathbf{B}_L^M$ denote the normalized ψ_L^M. We obtain $\mathbf{B}_2^1(1, 1)$ from (54) by dividing through by the square root of the sum of the squares of the coefficients giving

$$\mathbf{B}_2^1 = \frac{1}{\sqrt{2}}\mathbf{H}_1^1\mathbf{K}_1^0 + \frac{1}{\sqrt{2}}\mathbf{H}_1^0\mathbf{K}_1^1 \tag{55}$$

In a similar manner we find

$$\mathbf{B}_1^1 = -\frac{1}{\sqrt{2}}\mathbf{H}_1^1\mathbf{K}_1^0 + \frac{1}{\sqrt{2}}\mathbf{H}_1^0\mathbf{K}_1^1 . \tag{56}$$

Because we now have an orthonormal system expressed in terms of another orthonormal system the matrix of the coefficients is unitary, and, being real, is orthogonal. By the same order of steps the normalized functions $\mathbf{B}_L^0(1, 1)$ are, from (50),

	$\mathbf{H}_1^1\mathbf{K}_1^{-1}$	$\mathbf{H}_1^0\mathbf{K}_1^0$	$\mathbf{H}_1^{-1}\mathbf{K}_1^1$
$\mathbf{B}_2^0(1, 1)$	$\frac{1}{\sqrt{6}}$	$\frac{2}{\sqrt{6}}$	$\frac{1}{\sqrt{6}}$
$\mathbf{B}_1^0(1, 1)$	$-\frac{1}{\sqrt{2}}$	0	$\frac{1}{\sqrt{2}}$
$\mathbf{B}_0^0(1, 1)$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$

(57)

The equations for the $\mathbf{B}_L^{-1}(1, 1)$ will evidently yield the same matrix of coefficients as the $\mathbf{B}_L^1(1, 1)$ given by (55) and (56). We shall always have

$$\mathbf{B}_{l+l'}^{l+l'}(l, l') = \mathbf{H}_l^l\mathbf{K}_{l'}^{l'}; \qquad \mathbf{B}_{l+l'}^{-l-l'}(l, l') = \mathbf{H}_l^{-l}\mathbf{K}_{l'}^{-l'} \tag{58}$$

as exemplified by the first and last rows of (50): normalization cannot alter the single element $+1$. The $\mathbf{B}_L^M(1, 1)$ are thus completely determined.

Useful however are the normalization factors for the ψ_L^M, namely the sum of the squares of the coefficients in (54) and like cases, for this sum is equal to the integral of the square of the absolute value of ψ_L^M taken (for both sets of variables) over the surface of the unit sphere. Hence in working out (55)–(57) a record was kept of these sums, which may be

denoted by (ψ_L^M, ψ_L^M). Omitting the common factor $16\pi^2$ these normalizing factors are

$$\begin{array}{lcccccc} \text{For} & \psi_2^2(1,1) & \psi_2^1(1,1) & \psi_1^1(1,1) & \psi_2^0(1,1) & \psi_1^0(1,1) & \psi_0^0(1,1) \\ (\psi,\psi) = & \frac{4}{9} & \frac{1}{9} & \frac{4}{9} & \frac{2}{27} & \frac{2}{9} & \frac{4}{3} \end{array} \tag{59}$$

It is unnecessary to tabulate for negative M since the factor is always the same as for corresponding positive M.

Nor is it in general necessary, in expressing the $\psi_L^M(l, l')$ in terms of the $H_l^m K_{l'}^{m'}$, or the $\mathbf{B}_L^M$ in terms of the $\mathbf{H}_l^m \mathbf{K}_{l'}^{m'}$ to write out the equations for negative M. For we can always obtain ψ_L^{-M} from ψ_L^M by, first, changing m and m' into $-m$ and $-m'$ and, second, multiplying by $(-1)^k$ where $k = l + l' - L$; this relation is easily deduced from (39) and is illustrated by the expansion of ψ_1^{-1} as given by (50).

While it is true that orthogonal matrices like (57) are useful in some applications, yet if we have occasion to write out the $\mathbf{B}_L^M$ explicitly in terms of the variables x, y, z, x_1, y_1, z_1 it is more expeditious first to write out the ψ_L^M by equations like (50) and afterwards to normalize, from a table like (59). Suppose, for example, we wish to obtain $\mathbf{B}_0^0(1, 1)$. We first write from (50)

$$\psi_0^0(1, 1) = H_1^1 K_1^{-1} - 2H_1^0 K_1^0 + H_1^{-1} K_1^1. \tag{60}$$

We then have from (20) and (21)

$$\begin{array}{lll} H_1^1 = x + iy, & H_1^0 = z, & H_1^{-1} = -(x - iy); \\ K_1^1 = x_1 + iy_1, & K_1^0 = z_1, & K_1^{-1} = -(x_1 - iy_1). \end{array} \tag{61}$$

Hence

$$\begin{aligned} \psi_0^0 &= -(x + iy)(x_1 - iy_1) - 2zz_1 - (x - iy)(x_1 + iy_1) \\ &= -2(xx_1 + yy_1 + zz_1) \end{aligned} \tag{62}$$

which is an absolute invariant under rotation, as every $\psi_0^0(l, l)$ must be. Obtaining the proper normalizing factor from (59) we then have

$$\mathbf{B}_0^0(1, 1) = \frac{\sqrt{3}}{4\pi}(xx_1 + yy_1 + zz_1) \tag{63}$$

which is normalized with reference to the surface of the unit sphere having centre at the origin.

Similarly we find

$$\begin{aligned}\psi_1^1(1,1) &= -(x+iy)z_1 + z(x_1+iy_1) = zx_1 - xz_1 + i(zy_1 - yz_1)\\ &= -i[yz_1 - zy_1 + i(zx_1 - xz_1)]\end{aligned} \tag{64}$$

which, if we adopt the notation

$$X = yz_1 - zy_1, \qquad Y = zx_1 - xz_1, \qquad Z = xy_1 - yx_1 \tag{65}$$

becomes

$$\psi_1^1(1,1) = -i[X + iY]. \tag{66}$$

Similarly we find

$$\psi_1^0(1,1) = -iZ, \qquad \psi_1^{-1}(1,1) = +i[X - iY] \tag{67}$$

whence, comparing with (61), we see that ψ_1^1, ψ_1^0, and ψ_1^{-1} actually do transform like H_1^1, H_1^0, and H_1^{-1}, for X, Y, and Z transform like x, y, z as is well known. The $\mathbf{B}_1^M(1,1)$ may now be found by aid of (59).

The $\psi_2^M(1,1)$ are similarly found to be

$$\begin{aligned}\psi_2^2(1,1) &= (x+iy)(x_1+iy_1);\\ \psi_2^1(1,1) &= \tfrac{1}{2}[(x+iy)z_1 + z(x_1+iy_1)];\\ \psi_2^0 &= zz_1 - \tfrac{1}{3}(xx_1 + yy_1 + zz_1);\\ \psi_2^{-1} &= -\tfrac{1}{2}[(x-iy)z_1 + z(x_1 - iy_1)];\\ \psi_2^{-2} &= (x-iy)(x_1 - iy_1)\end{aligned} \tag{68}$$

which when compared with

$$\begin{aligned}H_2^2 = (x+iy)^2; \qquad H_2^1 = (x+iy)z; \qquad H_2^0 = z^2 - \tfrac{1}{3}r^2;\\ H_2^{-1} = -(x-iy)z; \qquad H_2^{-2} = (x-iy)^2\end{aligned} \tag{69}$$

make it clear that the ψ_2^M actually do transform under rotation like the H_2^m, since the first set of functions reduces precisely to the second when we put $x_1 = x$, $y_1 = y$, $z_1 = z$.

As the final step in working out the case $l = l' = 1$ we now introduce the normalizing factors from (59), as has already been done for $\mathbf{B}_0^0$, and tabulate results:

$$\mathbf{B}_2^2(1,1) = \frac{3}{8\pi}(x+iy)(x_1+iy_1); \qquad \mathbf{B}_2^1 = \frac{3}{8\pi}[(x+iy)z_1 + z(x_1+iy_1)];$$

$$\mathbf{B}_2^0 = \frac{3\sqrt{6}}{8\pi}\left[zz_1 - \frac{1}{3}(xx_1+yy_1+zz_1)\right]$$

$$\mathbf{B}_1^1 = \frac{3}{8\pi}(X+iY) \qquad \mathbf{B}_1^0 = \frac{3\sqrt{2}}{8\pi}Z; \tag{70}$$

$$\mathbf{B}_0^0 = \frac{\sqrt{3}}{4\pi}(xx_1+yy_1+zz_1)$$

It is scarcely necessary to tabulate **B**'s with negative M for these are found by taking the conjugate and multiplying by $(-1)^M$.

6. **The case $l = 2, l' = 1$.** Having worked out in rather full detail the case $l = l' = 1$, in order to exemplify the method, it will be sufficient in other cases merely to tabulate the principal results. These are, first, the matrix which determines the **B**'s in terms of the $\mathbf{H}_l^m\mathbf{K}_m^{l'}$, and, second, the equations of type (70) which give the simplified **B**'s in a form such that their covariance under rotation (in the case $\mathbf{B}_0^0$ their invariance) is directly manifest.

For the case $l = 2, l' = 1$ we have, omitting for convenience the common denominator 30 from each element of the matrix,

	$\mathbf{H}_2^0\mathbf{K}_1^0$	$\mathbf{H}_2^{-1}\mathbf{K}_1^1$	$\mathbf{H}_2^2\mathbf{K}_1^{-1}$	$\mathbf{H}_2^1\mathbf{K}_1^0$	$\mathbf{H}_2^0\mathbf{K}_1^1$	$\mathbf{H}_2^1\mathbf{K}_1^{-1}$	$\mathbf{H}_2^0\mathbf{K}_1^0$	$\mathbf{H}_2^{-1}\mathbf{K}_1^1$
$\mathbf{B}_3^2$	$10\sqrt{3}$	$10\sqrt{6}$	·	·	·	·	·	·
$\mathbf{B}_2^2$	$-10\sqrt{6}$	$10\sqrt{3}$	·	·	·	·	·	·
$\mathbf{B}_3^1$	·	·	$2\sqrt{15}$	$4\sqrt{30}$	$6\sqrt{10}$	·	·	·
$\mathbf{B}_2^1$	·	·	$-10\sqrt{3}$	$-5\sqrt{6}$	$15\sqrt{2}$	·	·	·
$\mathbf{B}_1^1$	·	·	$6\sqrt{15}$	$-3\sqrt{30}$	$3\sqrt{10}$	·	·	·
$\mathbf{B}_3^0$	·	·	·	·	·	$6\sqrt{5}$	$6\sqrt{15}$	$6\sqrt{5}$
$\mathbf{B}_2^0$	·	·	·	·	·	$-15\sqrt{2}$	0	$15\sqrt{2}$
$\mathbf{B}_1^0$	·	·	·	·	·	$3\sqrt{30}$	$-6\sqrt{10}$	$3\sqrt{30}$

(71)

To save space, the equation $\mathbf{B}_3^3 = \mathbf{H}_2^2\mathbf{K}_1^1$ has been omitted, and, as pointed out above, tabulation for negative M is unnecessary. The orthogonality of the matrix (after inserting the common denominator 30 for every element) is readily verified.

In carrying out the computation a record was kept of the normalizing

factors. For the H_2^m we have N_2^2, N_2^1, and N_2^0 (as defined by (51)) equal to the common factor $\frac{4\pi}{45}$ multiplied by 24, 6, and 4, respectively. For the (ψ_L^M, ψ_L^M) we have

$$\begin{array}{lccccccccc} \text{For} & \psi_3^3 & \psi_3^2 & \psi_2^2 & \psi_3^1 & \psi_2^1 & \psi_1^1 & \psi_3^0 & \psi_2^0 & \psi_1^0 \\ (\psi,\psi) = & 240 & 40 & 135 & 16 & 45 & 400 & 12 & 30 & 200 \end{array} \tag{72}$$

each of these numbers to be multiplied by the common factor $\frac{16\pi^2}{675}$.

In tabulating the explicit values of the ψ_L^M it will be convenient to write

$$xx_1 + yy_1 + zz_1 = q. \tag{73}$$

The F-biharmonics ψ_3^M are, for the case $l = 2, l' = 1$,

$$\psi_3^3 = H_2^2 K_1^1 = (x + iy)^2(x_1 + iy_1).$$

$$\psi_3^2 = \frac{1}{3}H_2^2K_1^0 + \frac{2}{3}H_2^1K_1^1 = \frac{1}{3}(x + iy)^2z_1 + \frac{2}{3}(x + iy)z(x_1 + iy_1).$$

$$\psi_3^1 = \frac{1}{15}H_2^2K_1^{-1} + \frac{8}{15}H_2^1K_1^0 + \frac{2}{5}H_2^0K_1^1$$

$$= \frac{2}{3}(x + iy)\left(zz_1 - \frac{q}{5}\right) + \frac{1}{3}(x_1 + iy_1)\left(z^2 - \frac{1}{3}r^2\right).$$

$$\psi_3^0 = \frac{1}{5}H_2^1K_1^{-1} + \frac{3}{5}H_2^0K_1^0 + \frac{1}{5}H_2^{-1}K_1^1 = z^2z_1 - \frac{2}{5}zq - \frac{1}{5}z_1r^2. \tag{74}$$

and by comparison with

$$H_3^3 = (x + iy)^3, \quad H_3^2 = (x + iy)^2z, \quad H_3^1 = (x + iy)\left(z^2 - \frac{1}{5}r^2\right),$$
$$H_3^0 = z^3 - \frac{3}{5}r^2z, \tag{75}$$

we can verify that the ψ_3^M actually do transform under rotation like the H_3^m, noting that because the two sets of variables undergo the same rotation by hypothesis, q transforms like r^2, while z^2z_1 transforms like z^3, and so on.

Tht D-biharmonics $\psi_2^M(2, 1)$ are

$$\psi_2^2 = H_1^2K_1^1 - H_2^2K_1^0 = -i(x + iy)(X + iY).$$

$$\psi_2^1 = -\frac{1}{4}H_2^2K_1^{-1} - \frac{1}{2}H_2^1K_1^0 + \frac{3}{4}H_2^0K_1^1 = -\frac{i}{2}[(X + iY)z + (x + iy)Z].$$

$$\psi_2^0 = -\frac{1}{2}H_2^1K_1^{-1} + \frac{1}{2}H_2^{-1}K_1^1 = -izZ = -i\left[zZ - \frac{1}{3}(xX + yY + zZ)\right] \tag{76}$$

where the last step holds because $xX + yY + zZ = 0$. These results may be compared with (69).

Finally the P-biharmonics are for this case

$$\psi_1^1 = H_2^2K_1^{-1} - 2H_2^1K_1^0 + H_2^0K_1^1 = -2(x + iy)q + \frac{2}{3}(x_1 + iy_1)r^2.$$

$$\psi_1^0 = H_2^1K_1^{-1} - 2H_2^0K_1^0 + H_2^{-1}K_1^1 = -2zq + \frac{2}{3}z_1r^2. \tag{77}$$

In the above tabulations functions with negative M have been omitted as already explained. The corresponding $\mathbf{B}_L^M(2, 1)$ may be at once written down by aid of the table (72). A factor of absolute value unity, such as $-i$, or -1, may be omitted in the normalization.

7. **A second method of normalization.** In the two examples which have been worked through above, the factors for normalizing the ψ's were found as sums of squares of the coefficients when any ψ was expressed as a linear combination of the normalized $\mathbf{H}_l^m\mathbf{K}_{l'}^{m'}$. However, when the normalizing factor for any ψ_L^L has been thus found, that for ψ_L^M can be at once obtained by dividing by the binomial coefficient $\binom{2L}{L-M}$. This is seen by inspection of (51), noting that, since the ψ_L^M transform like the H_L^M, the normalizing factors must be proportional. This relation serves as a good check on the computations, and may be verified on (59) and (72). As above remarked, a factor of absolute value unity may be neglected in normalization, but if such a factor is cancelled from any ψ_L^M it should be likewise cancelled from all other functions of the same primitive manifold in order to preserve the isomorphism with the $\mathbf{H}_L^M$.

8. **A general formula for the coefficients.** When the values of l and of l' are numerically assigned, the most rapid way of calculating the $\psi_L^M(l, l')$ is undoubtedly that illustrated in the above examples. It is also possible however to write a general formula for the elements of the

matrices of the type of (50), and such an expression is not without interest. Consider the fact that the ψ_L^M were obtained from (39) by successive operation with $(D_{\eta} + D_{\eta_1})$. Let now this operator be applied t times to the general product $H_l^{l-j} K_{l'}^{l'-j'}$, where j and j' are integers not greater than $2l$ and $2l'$ respectively. By a rather easy induction the result is

$$(D_{\eta} + D_{\eta_1})^t H_l^{l-j} K_{l'}^{l'-j'} = t!\left\{\binom{2l-j}{t} H_l^{l-j-t} K_{l'}^{l'-j'} + \binom{2l-j}{t-1}\binom{2l'-j'}{1} H_l^{l-j-t-1} K_{l'}^{l'-j'-1} + \cdots + \binom{2l-j}{t-w}\binom{2l'-j'}{w} H_l^{l-j-t+w} K_{l'}^{l'-j'-w} + \cdots\right\}$$

$$= t! \sum_{w=0}^{t} \binom{2l-j}{t-w}\binom{2l'-j'}{w} H_l^{l-j-t+w} K_{l'}^{l'-j'-w} \qquad (78)$$

If we apply this result to the successive terms on the right of (39) and collect like terms we find

$$\psi_L^{L-t}\binom{2L}{t} = \sum_{w=0}^{k+t} \sum_{s=0}^{k} (-1)^s \binom{k}{s}\binom{2l-k+s}{t-w+s}\binom{2l'-s}{w-s} H_l^{l-k-t+w} K_{l'}^{l'-w} \qquad (79)$$

where the summation is over all values of s and of w which have meaning: if the lower index of any binomial coefficient is negative the term containing this binomial coefficient is non-existent; if the lower index of any binomial coefficient is zero then this binomial coefficient is $+1$. The summation on s gives the numerator of the general matrix element of type (50), while the common denominator for the row of elements determined by fixed given values of L and t is $\binom{2L}{t}$. The value of w determines the column of the matrix when the work is arranged as in (50). For convenience of reference the relations among the various indices may be tabulated as follows,

$$M = L - t = l + l' - k - t, \qquad (80)$$

$$m = l - k - t + w, \qquad m' = l' - w \qquad (81)$$

$$t = l + l' - m - m' - k, \quad w = l' - m', \quad t - w = l - m - k. \qquad (82)$$

As in (37), k numbers the rows of the trapezoidal matrix (32), running from 0 to $2l'$.

Hence the coefficient of $H_l^m K_{l'}^{m'}$ in (79) may be written

$$\sum_{s=0}^{k} (-1)^s \binom{k}{s} \binom{2l-k+s}{l-m-k+s} \binom{2l'-s}{l'-m'-s} \div \binom{2L}{t} \tag{83}$$

a function of binomial coefficients. The simplicity of this result is a consequence of choosing the H_l^m as basis functions in the manifold of ordinary spherical harmonics.

As an illustration, let us verify the coefficient -2 of $H_2^0 K_1^0$ in the expression (77) for $\psi_1^0(2, 1)$. Putting $l = 2$, $l' = 1$, $k = 2$, $t = 1$ we have by (83)

$$\sum_{s=0}^{2} (-1)^s \binom{2}{s} \binom{2+s}{s} \binom{2-s}{1-s} \div \binom{2}{1} = \left[\binom{2}{1} - \binom{2}{1}\binom{3}{1}\right] \div \binom{2}{1} = -2. \tag{84}$$

The term with $s = 2$ is to be taken as zero, since there is no $\binom{0}{-1}$.

As another illustration let us verify the coefficient of $H_2^1 K_1^0$ in the expression (74) for $\psi_3^1(2, 1)$, namely $\frac{8}{15}$. Here $k = 0$, $t = 2$, giving

$$\binom{0}{0}\binom{4}{1}\binom{2}{1} \div \binom{6}{2} = \frac{8}{15} \tag{85}$$

where $\binom{0}{0}$ is $+1$ according to the rule given above.

By the general formula (79) we may systematically develop the elements of the trapezoidal matrix (32) as functions of l and l'. By virtue of the lemma this is most conveniently done by columns. The first column is always $\psi_{l+l'}^{l+l'} = H_l^l K_{l'}^{l'}$. A compact method of tabulation may be illustrated by representing the second column as follows:

	$H_l^l K_{l'}^{l'-1}$	$H_l^{l-1} K_{l'}^{l'}$	*denominator*
$k = 0,\ t = 1;\ \psi_{l+l'}^{l+l'-1}$	$2l'$	$2l$	$2l + 2l'$
$k = 1,\ t = 0;\ \psi_{l+l'-1}^{l+l'-1}$	-1	$+1$	$+1$

(86)

	$H_l^l K_{l'}^{l'-2}$	$H_l^{l-1} K_{l'}^{l'-1}$	$H_l^{l-2} K_{l'}^{l'}$	*denominator*
$k=0,\ t=2$: $\psi_{l+l'-2}^{l+l'-2}$	$2l'(2l'-1)$	$8ll'$	$2l(2l-1)$	$(2l+2l')(2l+2l'-1)$
$k=1,\ t=1$: $\psi_{l+l'-3}^{l+l'-3}$	$-(2l'-1)$	$2(l'-l)$	$2l-1$	$2l+2l'-2$
$k=2,\ t=0$: $\psi_{l+l'-4}^{l+l'-4}$	1	-2	1	1

(87)

	$H_l^l K_{l'}^{l'-3}$	$H_l^{l-1} K_{l'}^{l'-2}$	$H_l^{l-2} K_{l'}^{l'-1}$	$H_l^{l-3} K_{l'}^{l'}$
$\psi_{l+l'-3}^{l+l'-3}$	$2l'(2l'-1)(2l'-2)$	$12ll'(2l'-1)$	$12ll'(2l-1)$	$2l(2l-1)(2l-2)$
$\psi_{l+l'-4}^{l+l'-3}$	$(2l'-1)(2l'-2)$	$(2l'-4l)(2l'-1)$	$(4l'-2l)(2l-1)$	$(2l-1)(2l-2)$
$\psi_{l+l'-5}^{l+l'-3}$	$2l'-2$	$2l-4l'+2$	$2l'-4l+2$	$2l-2$
$\psi_{l+l'-6}^{l+l'-3}$	-1	$+3$	-3	$+1$

(88)

where the numerators of the coefficients are arranged in matrix form, while the denominators, which are common to the elements of a row, are given at the right. The table is made from (79) by using the indicated values of k and t. We may compare with (50) when $l = l' = 1$, and with (74) and (76) with $l = 2, l' = 1$.

For the third and fourth columns of (32) we find results as tabulated in (87) and (88).

The fourth column of (32) is particularly interesting, since the number of elements in this column will be two, three, or four, according to the special case. (88) is the matrix of numerators only; the denominators are

for $\psi_{l+l'}^{l+l'-3}$: $(2l + 2l')(2l + 2l' - 1)(2l + 2l' - 2)$;

for $\psi_{l+l'-1}^{l+l'-3}$: $(2l + 2l' - 2)(2l + 2l' - 3)$;

for $\psi_{l+l'-2}^{l+l'-3}$: $2l + 2l' - 4$; and $+1$ for $\psi_{l+l'-3}^{l+l'-3}$.

In the case $l = l' = 1$ the last row does not exist since L is never negative; and the third row also falls out, since ψ_0^{-1} is meaningless. The first two rows give a result in agreement with (50). In the case $l = 2, l' = 1$ the last row falls away because there is no ψ_0^0 unless $l = l'$. The other rows give results in agreement with Art. 6, the elements of the first column being zero.

In the cases just tabulated each numerator is a product of linear factors in l and l' with rational coefficients. In general such is not the case. If we work out the fifth column of (32) we find that the coefficient of $H_l^{l-2} K_{l'}^{l'-2}$ in the expansion of $\psi_{l+l'-2}^{l+l'-4}$, found from (79) by taking $k = t = w = 2$, contains the factor

$$2l^2 - 8ll' + 2l'^2 + 3l + 3l' - 2 \tag{89}$$

which is not factorable with rational coefficients.

9. **A general formula for normalization.** On the other hand the normalization factor (ψ_L^M, ψ_L^M) can be thrown into a very simple form in which numerator and denominator are products of linear factors. Consider first the cases (ψ_L^L, ψ_L^L). By (39) and (51)

$$(\psi_L^L, \psi_L^L) = N_l^l N_{l'}^{l'} \sum_{s=0}^{k} \frac{\binom{k}{s}^2}{\binom{2l}{k-s}\binom{2l'}{s}} \tag{90}$$

It will now be shown that the summation on s is equal to the fraction

$$\frac{\binom{2l+2l'-k+1}{k}}{\binom{2l}{k}\binom{2l'}{k}} \tag{91}$$

To prove this I shall first prove the more general identity

$$\begin{aligned}
&(a+x)(a+2x)(a+3x)\cdots(a+nx)\\
&\quad + n(a+x)(a+2x)(a+3x)\cdots(a+nx-x)\cdot(b+x) +\\
&\quad \frac{n(n-1)}{2}(a+x)(a+2x)(a+3x)\cdots(a+nx-2x)(b+x)(b+2x)\\
&\quad + \cdots + \binom{n}{s}(a+x)(a+2x)(a+3x)\cdots\\
&\quad (a+nx-sx)\cdot(b+x)(b+2x)(b+3x)\cdots(b+sx)+\cdots\\
&\quad \equiv (a+b+nx+x)(a+b+nx)(a+b+nx-x)\cdots(a+b+2x)
\end{aligned} \tag{92}$$

the right side being the product of n factors each of which exceeds its successor by x. It is assumed that n is a positive integer, but a, b and x may be any quantities of ordinary algebra. We now multiply both sides by $a+b+nx+2x$. On the left side arrange in the following way: the first term is multiplied by $(a+nx+x)+(b+x)$; the second term is multiplied by $(a+nx)+(b+2x)$; the term with coefficient $\binom{n}{s-1}$ is multiplied by $(a+\{n+2-s\}\,x)+(b+sx)$; the term with coefficient $\binom{n}{s}$ is multiplied by $(a+\{n+1-s\}\,x)+(b+sx+x)$; and so on. Collecting terms, the coefficient of

$$(a+x)\,(a+2x)\cdots(a+\{n+1-s\}\,x)\,(b+x)\,(b+2x)\cdots(b+sx)$$

is $\binom{n}{s}+\binom{n}{s-1}$ which equals $\binom{n+1}{s}$. Hence if the identity holds for an integer n it holds for $n+1$. But

$$(a+x)+(b+x)=a+b+x+x,$$

that is, the identity holds for $n=1$; hence it holds for all integers n.

It is interesting to notice that the identity (92) reduces to the binomial theorem for positive integral n if we let x equal zero.

Now if we equate the summation of (90) to the required expression (91) and cancel factors common to numerator and denominator of the various terms we find

$$\begin{aligned}
&2l\,(2l-1)\,(2l-2)\cdots(2l-k+1)\\
&\quad+k(2l-1)\,(2l-2)\cdots(2l-k+1)\cdot(2l'-k+1)+\cdots\\
&\quad+\binom{k}{s}(2l-s)(2l-s-1)\cdots(2l-k+1)\cdot\\
&\quad(2l'-k+1)\,(2l'-k+2)\cdots(2l'-k+s)+\cdots\\
&\quad=(2l+2l'-k+1)\,(2l+2l'-k)\cdots(2l+2l'-2k+2) \qquad (93)
\end{aligned}$$

which is a special case of (92) obtained by letting $x = +1$, $n = k$, $a = 2l - k$, $b = 2l' - k$. By using Art. 7 we have for the generalised (ψ, ψ),

$$\binom{2l}{k}\binom{2l'}{k}\binom{2L}{L-M}(\psi_L^M, \psi_L^M) = N_l^l N_{l'}^{l'}\binom{2l+2l'-k+1}{k} \qquad (94)$$

where N_l^l is the first factor on the right of (51).

The summation on s in (79) evidently gives an integer, easily calculated in any special case, since it is a function of binomial coefficients only. If we let this integer be denoted by I_w we can collect results and obtain a very simple general formula for the normalized function **B** in terms of the normalized **HK**. For, first normalizing the **HK** we have by (79) and (51) and (52),

$$\psi_L^{L-t}\binom{2L}{t} = \sum_{w=0}^{k+t} I_w \mathbf{H}_l^{l-k-t+w}\mathbf{K}_{l'}^{l'-w}\frac{(N_l^l N_{l'}^{l'})^{\frac{1}{2}}}{\left[\binom{2l}{k+t-w}\binom{2l'}{w}\right]^{\frac{1}{2}}} \qquad (95)$$

Normalizing ψ by aid of (94) then gives, noting that $L - M = t$,

$$\mathbf{B}_L^{L-t} = \left[\frac{\binom{2l}{k}\binom{2l'}{k}}{\binom{2L}{t}\binom{2l+2l'-k+1}{k}}\right]^{\frac{1}{2}}\sum_{w=0}^{k+t} I_w \frac{\mathbf{H}_l^{l-k-t+w}\mathbf{K}_{l'}^{l'-w}}{\left[\binom{2l}{k+t-w}\binom{2l'}{w}\right]^{\frac{1}{2}}} \qquad (96)$$

where the coefficients are functions of binomial coefficients only, since

$$I_w = \sum_{s=0}^{k}(-1)^s\binom{k}{s}\binom{2l-k+s}{t-w+s}\binom{2l'-s}{w-s} \qquad (97)$$

The results (71), (55), (56) and (57), which were obtained by direct computation, may of course be checked by (96). The formulas (80), (81), (82) are useful in comparing indices.

For example, to obtain the coefficient of $\mathbf{H}_2^1\mathbf{K}_1^0$ in the development of $\mathbf{B}_2^1(2, 1)$ we put $l = 2$, $l' = 1$, $k = 1$, $t = 1$, because we are dealing with the second element of the second row of the matrix of form (32), and also put $w = 1$ to make $m = 1$ and $m' = 0$. From (97) we find $l_1 = -2$ and from (96) obtain $-\frac{5\sqrt{6}}{30}$ agreeing with this coefficient as tabulated in (71).

10. **Application to the expansion of a product of two ordinary spherical harmonics.** There are a number of problems to which the biharmonics $\mathbf{B}_L^M(l, l')$ stand in close relation. Some of these will be very briefly indicated. If, for example, we let the two sets of variables x, y, z and x_1, y_1, z_1 become identical, i.e. let $x_1 = x$, $y_1 = y$, $z_1 = z$, then $\mathbf{K}_{l'}^{m'}$ becomes $\mathbf{H}_{l'}^{m'}$ and the general formula (96) yields relations between the products $\mathbf{H}_l^m\mathbf{H}_{l'}^{m'}$ and the values taken on by the $\mathbf{B}$'s under the same circumstances. But since the $\mathbf{B}$'s of any primitive manifold, i.e. for any fixed value of L, transform under rotation like the corresponding $\mathbf{H}$'s, each $\mathbf{B}_L^M$ must, when the two sets of variables become the same set, either vanish or become a power of r^2 times a constant multiple of $\mathbf{H}_L^M$. Since the product $\mathbf{H}_l^m\mathbf{K}_{l'}^{m'}$ is a polynomial homogeneous of degree $l + l'$ we may write, by an expansion of the same type as (5),

$$\mathbf{H}_l^m\mathbf{H}_{l'}^{m'} = Q_0\mathbf{H}_{l+l'}^{m+m'} + Q_1r^2\mathbf{H}_{l+l'-2}^{m+m'} + \cdots + Q_nr^{2n}\mathbf{H}_{l+l'-2n}^{m+m'} + \cdots, \quad (98)$$

where the Q_n are not functions of x, y, z but depend on l, l', m, m'. We are to obtain (98) by solving (96) for the required $\mathbf{H}_l^m\mathbf{K}_{l'}^{m'}$ and then letting $x_1 = x$ etc. It is obvious that any $\mathbf{B}$ for which k $(= l + l' - L)$ is odd must vanish when the two sets of variables coincide because the terms which would result are absent from (98). This is exemplified by the expressions for $\mathbf{B}_1^1$ and $\mathbf{B}_1^0$ in (70), and by (76), since X, Y, and Z vanish when $x = x_1$, $y = y_1$, $z = z_1$.

Solving the equations (96) is a simple matter because any matrix of the type (71) is orthogonal, so that solution consists merely in the interchange of rows with columns. This interchange is equivalent to shifting from summation on w to summation on k, in part exemplified above in obtaining (83). We write

$$I_k = \sum_{s=0}^{k} (-1)^s \binom{k}{s}\binom{2l-k+s}{l-m-k+s}\binom{2l'-s}{l'-m'-s} \quad (99)$$

which is the same as I_w except for the change of notation by (82). Similarly we express all indices in (96) in terms of l, l', m, m' and k and solve by changing the summation to k. Hence

$$\mathbf{H}_l^m \mathbf{K}_{l'}^{m'} = \frac{1}{\left[\binom{2l}{l-m}\binom{2l'}{l'-m'}\right]^{\frac{1}{2}}} \sum_{k=0}^{l+l'-m-m'} I_k \mathbf{B}_{l+l'-k}^{m+m'} \left[\frac{\binom{2l}{k}\binom{2l'}{k}}{\binom{2l+2l'-2k}{l+l'-m-m'-k}\binom{2l+2l'-k+1}{k}}\right]^{\frac{1}{2}} \tag{100}$$

which is the solution of (96).

To obtain (98) we have now only to find what the $\mathbf{B}$'s become when k is even and we let $x_1 = x$ etc. This also is not difficult, thanks to the identity (92). Let us revert once more to the biharmonics $\psi_{l+l'-k}^{l+l'-k}$ which stand at the left of their respective rows in the trapezoidal matrix (32). When $x_1 = x$ etc. and k is even these must reduce to r^k times a constant multiple of H_L^L, i.e. a constant multiple of $(x+iy)^L$. Substituting from (20) into (39), if Lim ψ denotes what ψ becomes when $x_1 = x$ etc.,

$$\operatorname{Lim} \psi_{l+l'-k}^{l+l'-k} = (x+iy)^L \sum_{s=0}^{k} (-1)^s \binom{k}{s} Z_l^{l-k+s} Z_{l'}^{l'-s} \tag{101}$$

in which the summation must reduce to a constant multiple of r^k and our problem is to find this constant. Inspection of (21) shows that those terms for which s is odd contribute nothing because Z_l^m contains z as a factor when $l-m$ is odd; and if $l-m = 2c$ the last term in the expansion of Z_l^m by (21) is

$$(-1)^c \frac{(2c)!}{D(l,c)} r^{2c}. \tag{102}$$

If, therefore, we write $k = 2u$ and $s = 2v$ the coefficient of r^k in the expansion of $Z_l^{l-k+s} Z_{l'}^{l'-s}$ by (21) is

$$(-1)^u \frac{(2u-2v)!}{D(l,u-v)} \cdot \frac{(2v)!}{D(l',v)} \tag{103}$$

whence from (101) the desired constant is

$$(-1)^u \sum_{v=0}^{u} \binom{2u}{2v} \frac{(2u-2v)!\,(2v)!}{D(l,u-v) \cdot D(l',v)} \tag{104}$$

By (12) this may be written

$$(-1)^u \sum_{v=0}^{u} \frac{(2u)!}{(2u-2v)!\,(2v)!} \cdot \frac{(2u-2v)!\,(2v)!}{2^{u-v}(u-v)!\,2^v v!} \cdot$$

$$\frac{1}{(2l-1)\cdots(2l-2u+2v+1)\cdot(2l'-1)\cdots(2l'-2v+1)}$$

wherein, however, we are to write $+1$ instead of $(2l'+1)$ and $(2l+1)$ when $v=0$ and $v=u$, respectively. This simplifies to

$$(-1)^u \frac{(2u)!}{2^u u!} \sum_{v=0}^{u} \binom{u}{v}$$

$$\cdot \frac{1}{(2l-1)\cdots(2l-2u+2v+1)\cdot(2l'-1)\cdots(2l'-2v+1)}$$

With the same replacement of $(2l'+1)$ and $(2l+1)$ by unity; and this again may be written as the product of a coefficient independent of v,

$$(-1)^u \frac{(2u)!}{2^u u!} \cdot \tag{105}$$

$$\frac{1}{(2l-1)(2l-3)\cdots(2l-2u+1)\cdot(2l'-1)(2l'-3)\cdots(2l'-2u+1)}$$

into the summation

$$\begin{aligned}&(2l'-1)\,(2l'-3)\cdots(2l'-2u+1)\\ &+u\,(2l'-3)\cdots(2l'-2u+1)\cdot(2l-2u+1)+\cdots\\ &+\binom{u}{v}(2l'-2v-1)\cdots(2l'-2u+1)\cdot(2l-2u+2v-1)\cdots\\ &(2l-2u+1)+\cdots+(2l-1)\cdots(2l-2u+1)\,.\end{aligned} \tag{106}$$

If in the left member of (92) we let $n=u$, and take $a+x=2l'-2u+1$, making $x=+2$; and similarly $b+x=2l-2u+1$, then we have the summation (106), which therefore equals

$$(2l+2l'-2u)\,(2l+2l'-2u-2)\cdots(2l+2l'-4u+2) \tag{107}$$

that is

$$\frac{2^u(l+l'-u)!}{(l+l'+2u)!} \tag{108}$$

By using the identity

$$\frac{(2u)!}{2^u u!} = 1.3.5. \cdots (2u-1) \tag{109}$$

and collecting results, we have for the required constant

$(-1)^u$

$$\frac{[1.3.5. \cdots (2u-1)]2^u[(l+l'-u)(l+l'-u-1)\cdots(l+l'-2u+1)]}{[(2l-1)(2l-3)\cdots(2l-2u+1)][(2l'-1)(2l'-3)\cdots(2l'-2u+1)]} \tag{110}$$

$$= N_u, \text{ say}.$$

That is

$$\text{Lim } \psi_{l+l'-2u}^{l+l'-2u} = N_u H_{l+l'-2u}^{l+l'-2u} r^{2u}. \tag{111}$$

Hence also

$$\text{Lim } \psi_{l+l'-2u}^{l+l'-2u-t} = N_u H_{l+l'-2u}^{l+l'-2u-t} r^{2u} \tag{112}$$

since all the ψ_L^M having a fixed value for L transform under rotation like the H_L^M having the same value of L and this will be true even when $x_1 = x$ etc. Hence the same factor N_u carries through the corresponding row of (32).

This result may be compared with those particular cases for which ψ has been given in explicit form. The case ψ_0^0 (1, 1) is given by (62). Here $l = l' = u = 1$ whence $N = -2$, while if $x_1 = x$ etc. (62) gives $-2r^2$, as it should, since $H_0^0(1, 1) = 1$.

Again, letting $l = 2$, $l' = 1$, $u = 1$, we obtain $N = -\frac{4}{3}$. From (77) we obtain

$$\text{Lim } \psi_1^1(2, 1) = -\frac{4}{3}(x + iy)r^2$$

which is correct, since $H_1^1 = x + iy$.

It may be remarked that N_0 is $+1$, no matter what l and l' may be, for

$$\psi_{l+l'}^{l+l'} = H_l^l K_{l'}^{l'} \tag{113}$$

by the definition (39). The formula (110) becomes meaningless for $u = 0$.

To obtain the coefficients Q_u of (98) we have now merely to normalize

(112) and substitute in (100). The normalization is by (94) and (51), giving

$$\operatorname{Lim} \mathbf{B}_L^M = \left[\frac{\binom{2l}{2u}\binom{2l'}{2u} N_L^L}{\binom{2l+2l'-2u+1}{2u} N_l^l N_{l'}^{l'}} \right]^{\frac{1}{2}} N_u r^{2u} \mathbf{H}_L^M , \qquad (114)$$

where $L = l + l' - 2u$ and $M = l + l' - 2u - t$; but it is interesting that t does not occur in the coefficient on the right, a result that might have been anticipated; for although the **B**'s of a particular primitive manifold, which are normalized with respect to two distinct sets of variables, do not remain normalized with respect to the one set when $x = x_1$ etc., they must reduce to a constant multiple of the corresponding **H**'s because first, similarity of transformation is preserved, and, second, normalization is invariant under rotation. We now substitute in (100) and find for the general coefficient Q_u of (98)

$$Q_u = \left[\frac{N_L^M}{N_l^m N_{l'}^{m'}} \right]^{\frac{1}{2}} \frac{\binom{2l}{2u}\binom{2l'}{2u}}{\binom{2l}{l-m}\binom{2l'}{l'-m'}\binom{2l+2l'-2u+1}{2u}} I_{2u} N_u \qquad (115)$$

where the normalizing factors N_l^m etc. are defined by (51), and I_{2u} by (99), while N_u is the factor (110). The first factor on the right is for normalization, hence if this be omitted and if we write

$$H_l^m H_{l'}^{m'} = H_{l+l'}^{m+m'} + q_1 r^2 H_{l+l'-2}^{m+m'} + \cdots + q_u r^{2u} H_{l+l'-2u}^{m+m'} + \cdots \qquad (116)$$

we shall have

$$q_u = \frac{\binom{2l}{2u}\binom{2l'}{2u}}{\binom{2l}{l-m}\binom{2l'}{l'-m'}\binom{2l+2l'+2u-1}{2u}} I_{2u} N_u \qquad (117)$$

as the general coefficient in (116).

11. **Application to surface harmonics and to Adams' formula.** Obviously if we let $r = 1$ either of the formulas (115) or (117) gives the means of expanding the product of any two surface harmonics as a sum of harmonics, the latter formula being the more convenient if we do not happen to be interested in normalization with respect to the surface.

In the special case of zonal harmonics, i.e. $m = m' = 0$, formula (117)

is the equivalent of a well-known expansion proved by Adams[4] for the product of two zonal harmonics as a sum, which he applies to the calculation of the integral of the product of three such harmonics. A comparison of (117) with Adams' work yields some interesting identities and shows that, when $m = m' = 0$, the coefficient I_{2u} resolves into linear factors, although the manner in which this happens is quite different from the factorization which has been shown to occur in the first four columns of (32). It is planned to include some details of these relations in a later publication.

12. **Application to the integral of the product of three harmonics.** If we multiply both sides of (98) by $\mathbf{H}_{l''}^{m''}$ and integrate over the surface of the sphere assuming

$$m'' = -m - m'; \qquad l'' = l + l' - 2u_1 \tag{118}$$

where u_1 is a particular value of u actually found on the right side, the result of the integration is

$$(-1)^{m+m'} Q_{u_1}, \tag{119}$$

the sign factor being due to the fact that the conjugate of $\mathbf{H}_l^m$ is $(-1)^m \mathbf{H}_l^{-m}$ by (22). If l'' and m'' do not satisfy the given conditions the integral is zero.

13. **Application to polynomial expansions.** It is easy to see that any polynomial in the six variables x, y, z, x_1, y_1, z_1 can be written as a linear combination of the $\psi_L^M(l, l')$ the coefficients being power series in r_1^2 and r_1^2. For, first, any given polynomial can be taken as a sum of polynomials each homogeneous in each of the two sets of variables. Then, by (19), each of these homogeneous polynomials can be developed as a linear function of biharmonics. Collecting results we shall have a sum

$$\sum f_{l,l'} \psi_{l,l'} \tag{120}$$

wherein each coefficient f is a finite power series in r^2 and r_1^2. Finally we can expand each $\psi_{l,l'}$ in terms of the $\psi_L^M(l, l')$, and the simplest general method for doing so is to use infinitesimal rotations in a manner strictly analogous to the use of $r^2 \Delta$ in Dougall's expansion. We set

$$\psi = \sum \psi^{M_b}, \qquad (M_b = l + l', l + l' - 1, \cdots, -l - l') \tag{121}$$

[4] Proc. Roy. Soc. London **27**, 1878, 63.

each ψ^{M_b} to be some linear combination of the elements of the corresponding column of the trapezoidal matrix (32). We further write as a temporary notation

$$d = iD_x + iD_{x_1}. \tag{122}$$

Now construct an operator $\mathbf{D}_b$ by forming the product $P(d - M_a)$ of all the operators $d - M_a$ obtained by letting M_a vary from $l + l'$ to $-l - l'$ omitting only a fixed chosen value M_b, and dividing by the product of all the factors $M_b - M_a$ omitting of course $M_b - M_b$. That is

$$\mathbf{D}_b = \frac{\mathbf{P}(d - M_a)}{\mathbf{P}(M_b - M_a)}, \qquad (M_b \text{ fixed}; M_a = l + l', \cdots, -l - l' \text{ omitting } M_b). \tag{123}$$

By (33) each operator $d - M_a$ annuls every element in the column of (32) corresponding to M_a but multiplies by $M_b - M_a$ all elements of the column corresponding to M_b. Hence

$$\mathbf{D}_b\psi = \psi^{M_b}. \tag{124}$$

Therefore also

$$\sum \mathbf{D}_b\psi = \psi, \quad (M_b = l + l', l + l' - 1, \cdots, -l - l'). \tag{125}$$

To illustrate, let $l = l' = 1$. $\mathbf{D}_b$ is interpreted to mean $\mathbf{D}_{M_b}$. By (123) we then have

$$\mathbf{D}_2 = \frac{(d-1)(d-0)(d+1)(d+2)}{(2-1)(2-0)(2+1)(2+2)} = \frac{1}{24}(d^4 + 2d^3 - d^2 - 2d) \tag{126}$$

$$\mathbf{D}_1 = \frac{(d-2)(d-0)(d+1)(d+2)}{(1-2)(1-0)(1+1)(1+2)} = -\frac{1}{6}(d^4 + d^3 - 4d^2 - 4d) \tag{127}$$

$$\mathbf{D}_0 = \frac{(d-2)(d-1)(d+1)(d+2)}{(0-2)(0-1)(0+1)(0+2)} = \frac{1}{4}(d^4 - 5d^2 + 4) \tag{128}$$

and similarly

$$\mathbf{D}_{-1} = -\frac{1}{6}(d^4 - d^3 - 4d^2 + 4d) \tag{129}$$

$$\mathbf{D}_{-2} = \frac{1}{24}(d^4 - 2d^3 - d^2 + 2d) \tag{130}$$

which agrees with (125) since, by addition,

$$\mathbf{D}_2 + \mathbf{D}_1 + \mathbf{D}_0 + \mathbf{D}_{-1} + \mathbf{D}_{-2} = 1. \tag{131}$$

To take a simple example, let $\psi = xx_1$. We wish to find the components of xx_1 in the five manifolds determined by the respective columns of the matrix (40). By ordinary differentiation, using (122) and (28) we find

$$\begin{aligned} d(xx_1) &= i(yx_1 + xy_1); \qquad d^2(xx_1) = 2(xx_1 - yy_1); \\ d^3(xx_1) &= 4i(yx_1 + xy_1); \qquad d^4(xx_1) = 8(xx_1 - yy_1). \end{aligned} \tag{132}$$

Hence by (124) we have

$$\begin{gathered} \psi^2 = \frac{1}{4}(x + iy)(x_1 + iy_1); \qquad \psi^1 = 0; \qquad \psi^0 = \frac{1}{2}(xx_1 + yy_1); \\ \psi^{-1} = 0; \qquad \psi^{-2} = \frac{1}{4}(x - iy)(x_1 - iy_1). \end{gathered} \tag{133}$$

Having thus resolved ψ into components according to the columns of (32) we have to find the components of each ψ^{M_b} in the primitive manifolds determined by the rows. We let ψ now denote, for brevity, a particular ψ^{M_b} and set

$$\psi = \sum \psi_{L_e}, \qquad (L_e = l + l', l + l' - 1, \cdots, M_b) \tag{134}$$

each ψ_{L_e} to be some constant multiple of a definite element of the column of (32) which corresponds to M_b. We write as a further temporary notation

$$\partial = (D_\xi + D_{\xi_1})(D_\eta + D_{\eta_1}). \tag{135}$$

and

$$R_e = (L_e + M_b)(L_e - M_b + 1) \tag{136}$$

so that, by (31), R_c is the root of ∂ corresponding to the element ψ_{L_c} in the column in question. The operator $\partial - R_c$ annuls this element but multiplies by $R_e - R_c$ the element ψ_{L_e} which lies in a different row corresponding to L_e where

$$R_e = (L_e + M_b)(L_e - M_b + 1). \tag{137}$$

Now construct an operator $\mathbf{G}_e$ by forming the product $\mathbf{P}(\partial - R_c)$ of all the operators $\partial - R_c$ obtained by letting L_c vary from $l + l'$ down to M_b, omitting only a chosen value L_e, and dividing by the product $\mathbf{P}(R_e - R_c)$ of all the factors $R_e - R_c$, omitting of course $R_e - R_c$; M_b

remains constant throughout, having always the value indicated by the particular column of (32) concerned. That is

$$\mathbf{G}_e = \frac{\mathbf{P}(\partial - R_c)}{\mathbf{P}(R_e - R_c)}, \qquad (L_e \text{ and } M_b \text{ fixed; } L_c = l + l', l + l' - 1, \cdots, M_b, \text{ omitting } L_e). \tag{138}$$

Then

$$\mathbf{G}_e \psi = \psi_{L_e} \tag{139}$$

where it is assumed that ψ lies in the manifold corresponding to that particular column of (32) given by M_b. With the same assumption therefore

$$\sum \mathbf{G}_e \psi = \psi, \qquad (L_e = l + l', l + l' - 1, \cdots, M_b). \tag{140}$$

To illustrate, let $l = l' = 1$ as before, and take $M_b = 0$, i.e. deal with the central column of the matrix (40). Interpreting $\mathbf{G}_e$ to mean $\mathbf{G}_{L_e}$ we first compute

$$\begin{aligned} R_2 &= (2 + 0)(2 - 0 + 1) = 6, \\ R_1 &= (1 + 0)(1 - 0 + 1) = 2, \\ R_0 &= (0 + 0)(0 - 0 + 1) = 0, \end{aligned} \tag{141}$$

then

$$\mathbf{G}_2 = \frac{(\partial - 2)(\partial - 0)}{(6 - 2)(6 - 0)} = \frac{1}{24}(\partial^2 - 2\partial) \tag{142}$$

$$\mathbf{G}_1 = \frac{(\partial - 6)(\partial - 0)}{(2 - 6)(2 - 0)} = -\frac{1}{8}(\partial^2 - 6\partial) \tag{143}$$

$$\mathbf{G}_0 = \frac{(\partial - 6)(\partial - 2)}{(0 - 6)(0 - 2)} = \frac{1}{12}(\partial^2 - 8\partial + 12) \tag{144}$$

which agrees with (140) since by addition

$$\mathbf{G}_2 + \mathbf{G}_1 + \mathbf{G}_0 = 1. \tag{145}$$

To take an example, it was found above that the component of the biharmonic xx_1 in the manifold determined by the central column of (40) was $\frac{1}{2}(xx_1 + yy_1)$. By ordinary differentiation using (135) and (29) we find

$$\partial\left[\frac{1}{2}(xx_1 + yy_1)\right] = xx_1 + yy_1 - 2zz_1;$$
$$\partial^2\left[\frac{1}{2}(xx_1 + yy_1)\right] = 6xx_1 + 6yy_1 - 12\,zz_1. \tag{146}$$

Hence by (139) we have

$$\psi_2 = -\frac{1}{2}\left(zz_1 - \frac{1}{3}q\right) = -\frac{1}{2}\psi_2^0(1, 1) \text{ by (68)}; \qquad \psi_1 = 0;$$
$$\psi_0 = \frac{1}{3}q = -\frac{1}{6}\psi_0^0(1, 1) \text{ by (62)}. \tag{147}$$

Collecting results we have

$$xx_1 = \frac{1}{4}(x + iy)(x_1 + iy_1) - \frac{1}{2}\left(zz_1 - \frac{1}{3}q\right) + \frac{1}{3}q + \frac{1}{4}(x - iy)(x_1 - iy_1), \tag{148}$$

where the four components which appear are seen by their form to belong in the manifolds in question; and the other components are zero.

To see that the above method is analogous with the Dougall expansion, we have merely to note that, with the notation of (8), C_k^n is a root of the operator $r^2\Delta$ corresponding to the operand r^kH_n. The operator H of (9) can be written

$$H = \frac{(r^2\Delta - C_{n-2}^2)(r^2\Delta - C_{n-4}^4)\cdots(r^2\Delta - C_{n-2j}^{2j})\cdots}{(0 - C_{n-2}^2)(0 - C_{n-4}^4)\cdots(0 - C_{n-2j}^{2j})\cdots} \tag{149}$$

and corresponds to the root zero, hence gives a harmonic. Similarly the successive terms of (11) can be obtained as components corresponding to roots C_{l-2}^2, C_{l-4}^4 etc. of $r^2\Delta$. It will be noted that, in using the operators temporarily called d and ∂ in the above way, these operators had no multiple roots in the manifolds to which they were applied. The same is true of $r^2\Delta$ if this operator is used to obtain Dougall's expansion.

14. **Summary.** Art. 1. Biharmonics are harmonic in two sets of variables. The method of Dougall for expanding a polynomial in terms of harmonics is extended to biharmonics. Art. 2–7. A method is given for computing basis functions in primitive manifolds of biharmonics, each set of variables being subject to the same arbitrary rotation, and

for normalizing these functions. Art. 8, 9. A general formula is given for the coefficients in these normalized basis functions. Art. 10–13. Application is made to expanding a product of harmonics as a sum; to the integral of a product of three harmonics; and to the development of an arbitrary polynomial in two sets of variables in terms of the primitive basis functions.

ON MINKOWSKI'S DEFINITION OF LENGTH AND AREA[1]

By Philip Franklin

A method of defining the length of a closed convex curve in a plane in terms of area, or the area of a closed convex surface in space in terms of volume was given by Minkowski.[2] He used it to prove certain inequalities for convex bodies. We show here that a slight modification of the definition is applicable to any curve or surface made up of regular pieces. This modified definition gives a simpler approach to these concepts than the prevalent methods of defining them.

1. *Plane Curves.* Consider a simple arc of a curve in the plane, having a parametric representation $x = f(t)$, $y = g(t)$ where each of these functions has a continuous second derivative in the interval $t_1 \leqq t \leqq t_2$, and $[f'(t)]^2 + [g'(t)]^2 \neq 0$.

Our fundamental calculation for the length is as follows. We draw a circle of radius ϵ about each point of the arc as center, and consider the set of points S_ϵ, each of which is in or on at least one of these circles. If $M(S_\epsilon)$ means the 2-dimensional measure of S_ϵ, the arc length L of our curve is defined as

$$L = \lim_{\epsilon \to 0} \frac{M(S_\epsilon)}{2\epsilon}. \tag{1}$$

For an arc of the kind here considered, we may give this definition a different form. The radius of curvature (defined in terms of limiting intersection of normals) has a minimum η_1. Again, since the arc is simple, if the normal at any part P cuts the arc in some other point Q, there will be a minimum distance PQ, $2\eta_2$. Let the distance between the end points be $2\eta_3$. Let us take $\epsilon < \eta_1, \eta_2, \eta_3$ and lay off a distance ϵ on the normal in both directions at each point of the curve, giving a set of points N_ϵ. If we adjoin to N_ϵ two semicircles C_1, C_2 of radius ϵ at the end points of the curve, and take any point outside of the set $N_\epsilon + C_1 + C_2$, the minimum distance from this point to the arc will exceed ϵ,

[1] Presented to the American Mathematical Society.

[2] Minkowski, Über die Begriffe Länge, Oberfläche und Volumen, Jahrsber. der Deutschen Mathematiker Vereinigung, vol. 9 (1901), p. 115.

while any point inside the set $N_\epsilon + C_1 + C_2$ will have a minimum distance to the arc less than ϵ. Thus

$$S_\epsilon = N_\epsilon + C_1 + C_2,$$

and

$$M(S_\epsilon) = M(N_\epsilon) + \pi\epsilon^2,$$

so that

$$\lim_{\epsilon \to 0} \frac{M(S_\epsilon)}{2\epsilon} = \lim_{\epsilon \to 0} \frac{M(N_\epsilon)}{2\epsilon}.$$

This gives the alternative definition

$$L = \lim_{\epsilon \to 0} \frac{M(N_\epsilon)}{2\epsilon}.$$

We shall now show that the limit on the right exists, and equals the familiar integral expression for arc length. We begin by introducing co-ordinates for points in N_ϵ. Through any such point R, we draw a normal to the curve, cutting the curve at a point P. If t is the parameter for P, we call the positive direction along the tangent to the curve that of increasing t. The direction along the normal obtained from this by a positive ninety-degree rotation is taken as positive. Then t and r, the distance along the normal from P to R may be used as co-ordinates for R.

If we put

$$X(t) = \frac{-g'(t)}{\sqrt{f'(t)^2 + g'(t)^2}}, \qquad Y(t) = \frac{f'(t)}{\sqrt{f'(t)^2 + g'(t)^2}},$$

we have

$$x = f(t) + rX(t), \qquad y = g(t) + rY(t)$$

so that the Jacobian

$$J\begin{pmatrix} x & y \\ t & r \end{pmatrix} = \begin{vmatrix} f' + rX' & g' + rY' \\ X & Y \end{vmatrix} = \sqrt{f'^2 + g'^2} + r\frac{f''g' - f'g''}{f'^2 + g'^2}.$$

If we introduce s and α by

$$x' = s' \cos\alpha, \qquad y' = s' \sin\alpha, \qquad \text{so that } s = \int_{t_1}^{t} \sqrt{f'^2 + g'^2}\, dt$$

this simplifies to

$$J\begin{pmatrix} x & y \\ t & r \end{pmatrix} = s' - r\alpha'.$$

To compute $M(N_\epsilon)$, we have

$$\int_{N_\epsilon}\int dx\,dy = \int_{N_\epsilon}\int J\begin{pmatrix} x & y \\ t & r \end{pmatrix} dt\,dr = \int_{t_1}^{t_2} dt \int_{-\epsilon}^{\epsilon} dr(s' - r\alpha')$$

$$= 2\epsilon \int_{t_1}^{t_2} s'\,dt.$$

Thus

$$L = \lim_{\epsilon \to 0} \int_{t_1}^{t_2} s'\,dt = \int_{t_1}^{t_2} \sqrt{f'^2 + g'^2}\,dt.$$

It is interesting to observe that, when ϵ is subjected to the given limitations, and S_ϵ is replaced by N_ϵ, no limiting process is necessary.

If we laid off the distance on the normal on one side only, giving N_p on the positive side, we should have

$$M(N_p) = \epsilon \int_{t_1}^{t_2} s'\,dt - \frac{\epsilon^2}{2} \int_{t_1}^{t_2} \alpha'\,dt,$$

so that

$$L = \lim \frac{M(N_p)}{\epsilon}.$$

We note for comparison with the formulas in space that

$$M(N_p) = \epsilon \int_{s_1}^{s_2} ds - \frac{\epsilon^2}{2} \int_{\alpha_1}^{\alpha_2} d\alpha = \epsilon \Delta s - \frac{\epsilon^2}{2} \Delta\alpha$$

or

$$= \epsilon \int_{s_1}^{s_2} ds - \frac{\epsilon^2}{2} \int_{s_1}^{s_2} \frac{ds}{\rho}.$$

2. *Discontinuities.* The argument just given applies if the given arc has a finite number of points where the derivatives have discontinuities. For in this case we still have

$$L = \lim_{\epsilon \to 0} \frac{M(S_\epsilon)}{2\epsilon} = \lim_{\epsilon \to 0} \frac{M(N_\epsilon)}{2\epsilon},$$

since if there are K such points, the numerators will differ from the sums for the separate parts by at most $4K\pi\epsilon^2$.

For example, the result holds for polygons.

That an infinite number of discontinuities can not be so simply disposed of is seen by the example

$$x = t, \qquad y = f(t),$$

where $0 \leqq t \leqq 1$ and we write t in the base 3, and $f(t)$ in the base 2, replacing 0, 2 by 0, 1 if t contains no digit 1, and otherwise replacing all digits after the first 1 by zero, leaving this 1 unchanged, and treating earlier digits as before. For this arc, it may be shown that

$$\lim \frac{M(S_\epsilon)}{2\epsilon} = 2,$$

the length obtained by inscribed polygons, while

$$\lim \frac{M(N_\epsilon)}{2\epsilon} = 1$$

the value given by the integral formula. Whether this is typical, and the first definition is comparable in scope to the polygon definition, and the second corresponds to the integral formula, is an interesting open question.

3. *Space Curves.* Consider next an arc of a space curve, defined by functions having a continuous second derivative. We have define S_ϵ as before, but with spheres of radius ϵ, and put

$$L = \lim_{\epsilon \to 0} \frac{M(S_\epsilon)}{\pi\epsilon^2},$$

where M here means 3-dimensional measure.

As in the plane, we may use N_ϵ, instead of S_ϵ, obtained by drawing circles of radius ϵ in the normal planes, since their measures differ by a term of the third order in ϵ.

To calculate $M(N_\epsilon)$ we introduce co-ordinates u, v, along the principal normal (l, m, n) and binormal (λ, μ, ν) respectively. Thus

$$x = f(t) + lu + \lambda v, \qquad y = g(t) + mu + \mu v, \qquad z = h(t) + nu + \nu v,$$

and

$$J\begin{pmatrix} x, y, z \\ t, u, v \end{pmatrix} = \begin{vmatrix} f' + l'u + \lambda'v & g' + m'u + \mu'v & h' + n'u + \nu'v \\ l & m & n \\ \lambda & \mu & \nu \end{vmatrix}$$

$$= \sum \alpha(f' + l'u + \lambda'v) = \sum \alpha f' + u \sum \alpha l' + v \sum \alpha\lambda'$$

where (α, β, γ) is the direction of the tangent. Putting

$$s = \int_{t_1}^{t} \sqrt{f'^2 + g'^2 + h'^2}\, dt,$$

and using the Frenet formulas,[3]

$$f' = \alpha s', \qquad l' = -\left(\frac{\alpha}{\rho} + \frac{\lambda}{\tau}\right) s' \qquad \text{and} \qquad \lambda' = \frac{l}{\tau} s'$$

we have

$$J = s' - u\frac{s'}{\rho}.$$

Thus, if C denotes a circle of radius ϵ,

$$N(N_\epsilon) = \int_{N_\epsilon} dx\, dy\, dt = \int dt \int_C J\, du\, dv = \pi\epsilon^2 \int_{t_1}^{t_2} s'\, dt,$$

so that

$$L = \lim \frac{M(N_\epsilon)}{\pi\epsilon^2} = \lim \int_{t_1}^{t_2} \sqrt{f'^2 + g'^2 + h'^2}\, dt.$$

If an area other than a circle were used, there would be a term in

$$\int \frac{s'}{\rho}\, dt = \int^{\bullet} d\alpha$$

multiplied by a moment of the area, of the third order in ϵ, so that it would not change the limit giving L.

4. *Surface Area.* For a portion of a regular surface, given parametrically by

$$x = f(u, v), \qquad y = g(u, v), \qquad z = h(u, v),$$

with continuous second partial derivatives, as the image of some area in the u, v plane, we may define surface area by

$$A = \lim_{\epsilon \to 0} \frac{M(S_\epsilon)}{2\epsilon},$$

where S_ϵ is the set of points in or on a sphere with radius ϵ, and center some point of the surface. M means 3-dimensional measure. If the boundary of the surface has finite length, L, and we define N_ϵ by taking

[3] Eisenhart, Differential Geometry, Boston 1909, p. 17.

distances ϵ on the normal, $M(N_\epsilon)$ will differ from $M(S_\epsilon)$ by at most $\pi L\epsilon^2$, when ϵ is sufficiently small, so that we have

$$A = \lim_{\epsilon \to 0} \frac{M(N_\epsilon)}{2\epsilon}.$$

To calculate N_ϵ, we introduce co-ordinates u, v, and r along the normal (X, Y, Z)

$$x = f(u, v) + r\,X\,, \quad y = g(u, v) + r\,Y\,, \quad z = h(u, v) + r\,Z\,.$$

From this we have for the Jacobian:

$$J\begin{pmatrix} x, & y, & z \\ u, & v, & r \end{pmatrix} = \begin{vmatrix} f_u + rX_u & g_u + rY_u & h_u + rZ_u \\ f_v + rX_v & g_v + rY_v & h_v + rZ_v \\ X & Y & Z \end{vmatrix}$$

using the notation and formulas of Eisenhart[4] where

$$E\,du^2 + 2F\,du\,dv + G\,dv^2, \ (H^2 = EG - F^2)$$

and

$$D\,du^2 + 2D'\,du\,dv + D''\,dv^2$$

are the first and second fundamental forms, we may deduce that

$$J = H - rH\left(\frac{1}{\rho_1} + \frac{1}{\rho_2}\right) + r^2 H \frac{1}{\rho_1}\frac{1}{\rho_2}$$

where ρ_1 and ρ_2 are the principal radii of curvature.
Thus

$$M(N_\epsilon) = \int dx\,dy\,dz = \int du\,dv \int_{-\epsilon}^{\epsilon} J\,dr$$

$$= 2\epsilon \int H\,du\,dv + \frac{2\epsilon^3}{3}\int \frac{H}{\rho_1\rho_2}\,du\,dv$$

and

$$A = \lim_{\epsilon \to 0} \frac{M(N_\epsilon)}{2\epsilon} = \int H\,du\,dv\,,$$

the usual integral formula.

[4] Eisenhart, l. c., p. 114 f.

If the length ϵ were laid off on one side only, we should have

$$M(N_p) = \epsilon \int H\, du\, dv - \frac{\epsilon^2}{2} \int H\left(\frac{1}{\rho_1} + \frac{1}{\rho_2}\right) du\, dv + \frac{\epsilon^3}{3} \int H \frac{1}{\rho_1} \frac{1}{\rho_2}\, du\, dv.$$

5. *Discontinuities.* The definition of arc length of a space curve is still applicable if there are a finite number of points of discontinuity. Similarly, the definition of surface area continues to apply if there are a finite number of points or rectifiable curves of discontinuity. For, discontinuities of this type will only introduce infinitesimals of such an order that they disappear in the limit. Thus our definitions apply to curves or surfaces made up of a finite number of regular pieces.

ON THE CLOSURE OF THE SET $\{e^{in\theta}\}$

By Morris L. Kales

Mr. N. Levinson has communicated the following theorem to N. Wiener:

The set $\{e^{i\lambda_n\theta}\}$ ($\lambda_n \neq 0$, $n = 0, 1, 2, \cdots$) is closed L_2 over any interval of length less than 2π provided that $\underline{\lim} \frac{\lambda_n}{n} = 1$.[1]

In the following I consider a much less general theorem but give a very elementary proof.

THEOREM: The set $\{e^{in\theta}\}$ ($n = 0, 1, 2, \cdots$) is closed L_2 over any finite set of intervals the sum of whose lengths is less than 2π.

Let $f(\theta)$ be continuous in $[a, b]$ $(-\pi < a < b < \pi)$. Then by Weierstrass' theorem we can find a polynomial $\sum_0^N a_n\theta^n$ such that

$$\left| f(\theta) - \sum_0^N a_n\theta^n \right| < \epsilon \tag{1}$$

where ϵ is an arbitrary positive number. On the other hand, on the unit circle of the complex plane of z we have:

$$\theta^n = (-i)^n (\log z)^n \tag{2}$$

If we make a cut along the negative real axis $\log z$ will be analytic in the resulting region. By a theorem of Runge, we may represent any function which is analytic in a simply connected region by a series of polynomials which converges uniformly to the function in the interior of the region. If we cut off from the unit circle a small arc which intersects the negative real axis, the remainder of the curve will be wholly in the interior of a region in which $\log z$ is analytic. Hence for any $n \leqq N$ we can find a finite sum of polynomials $\sum_{m=0}^{N_n} P_{nm}(z)$ such that

$$\left| (-i)^n a_n (\log z)^n - \sum_{m=0}^{N_n} P_{nm}(z) \right| < \frac{\epsilon}{N} \quad (z = e^{i\theta} : a \leqq \theta \leqq b) \tag{3}$$

[1] For theorems on the closure of complex exponential functions see Paley and Wiener, "Fourier Transforms," American Mathematical Society Colloquium Publications, p. 86 ff.

i.e.

$$\left| a_n\theta^n - \sum_{m=0}^{M_n} a_{nm}e^{im\theta} \right| < \frac{\epsilon}{N} \tag{4}$$

combining (4) and (1) we have:

$$\left| f(\theta) - \sum_{n=0}^{N} \sum_{m=0}^{M_n} a_{nm}e^{im\theta} \right| < 2\epsilon \tag{5}$$

Thus

$$\int_a^b \left| f(\theta) - \sum_{n=0}^{N} \sum_{m=0}^{M_n} a_{nm}e^{im\theta} \right|^2 d\theta < 4\epsilon^2(b-a) \tag{6}$$

Now if $g(\theta)$ is a function belonging to L_2 we can find a continuous function $f(\theta)$ such that:

$$\int_a^b |\, g(\theta) - f(\theta)\,|^2\, d\theta < \lambda \tag{7}$$

(where λ is an arbitrary positive number)

From (6) and (7) and the Minkowski inequality we see finally that if $g(\theta)$ is any function of L_2 we can find an N and coefficients b_n such that

$$\int_a^b \left| g(\theta) - \sum_{0}^{N} b_n e^{in\theta} \right|^2 d\theta < \epsilon \tag{8}$$

Thus the set $\{e^{in\theta}\}$ $(n = 0, 1, 2, \cdots)$ is closed L_2 over any interval interior to $(-\pi, \pi)$. To prove the theorem in the case of an interval $(a, a + 2\pi)$ we have merely to consider the function $\varphi(\theta) = f(a + \pi + \theta)$.

Remark: In the preceding what we have actually shown is that the set $\{e^{in\theta}\}$ $(n = 0, 1, 2, \cdots)$ is complete L_2. But since completeness and closure are equivalent in L_2 the theorem follows.

If instead of making a cut along the negative real axis, we make it along some other radial line we can easily show that if from any interval of length 2π we remove an arbitrarily small interval, the set $\{e^{in\theta}\}$ $(n = 0, 1, 2, \cdots)$ is closed in the remaining interval or intervals. Consequently it is closed over any finite set of intervals, the sum of whose lengths is less than 2π, lying in an interval of length 2π.

A SIMPLE NON-LINEAR INTEGRAL EQUATION

By Morris L. Kales

The following problem affords a simple and interesting application of the theory of Fourier transforms.

Does there exist a function $f(x)$ which satisfies the equation:

$$f(x) = \int_{-\infty}^{\infty} f(x - \xi)\, f(\xi)\, d\xi$$

We shall show that when $f(x)$ belongs to L_1 there is only the trivial solution $f(x) \equiv 0$, but that when $f(x)$ belongs to L_2 there are non-trivial solutions.

Suppose first that $f(x) \subset L_1$. Let

$$g(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-iux}\, dx \tag{1}$$

Then since the Fourier transform of the "faltung" of two functions belonging to L_1 is equal to $\sqrt{2\pi}$ times the product of their Fourier transforms we have:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-iux}\, dx \int_{-\infty}^{\infty} f(x - \xi)\, f(\xi)\, d\xi = \sqrt{2\pi}\, g^2(u) \tag{2}$$

Thus:

$$g(u) = \sqrt{2\pi}\, g^2(u) \tag{3}$$

Hence for every u

$$g(u) = 0 \quad \text{or} \quad \frac{1}{\sqrt{2\pi}} \tag{4}$$

But the Fourier transform of a function belonging to L_1 is uniformly continuous. Hence, $g(u)$ must be either identically 0 or identically $\frac{1}{\sqrt{2\pi}}$. But since, by the Riemann-Lebesgue theorem, $g(u)$ tends to zero with increasing u it follows that

$$g(u) \equiv 0 \tag{5}$$

Since the Fourier transform of $f(x)$ is identically zero, $f(x)$ must be zero almost everywhere. If $f(x) = 0$ almost everywhere then $\int_{-\infty}^{\infty} f(x-\xi)f(\xi)\,d\xi \equiv 0$. Hence $f(x) \equiv 0$

If $f(x) \subset L_2$ and has the Fourier transform $g(u)$ we can again show in the same way that for every u:

$$g(u) = 0 \quad \text{or} \quad \frac{1}{\sqrt{2\pi}} \tag{1'}$$

However, in this case $g(u)$ need not be continuous, and consequently every function $g(u)$ which belongs to L_2 and satisfies (1') has as Fourier transform a function $f(x)$ which is a solution of our problem. Thus if E is any set of points of finite measure, the most general solution is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_E \frac{e^{iux}}{\sqrt{2\pi}}\,du$$

Note that $f(x)$ is uniformly continuous, since it is the "faltung" of two functions belonging to L_2.

As an example, let us consider a sequence of non-overlapping intervals (a_ν, b_ν) such that $\sum_{-\infty}^{\infty} (b_\nu - a_\nu)$ is finite. Let $g_\nu(u) = \frac{1}{\sqrt{2\pi}}$ over (a_ν, b_ν) and zero elsewhere. Let $g(u) = \sum_{-\infty}^{\infty} g_\nu(u)$. Then $\sum_{-\infty}^{\infty} g_\nu(u)\, e^{iux}$ is boundedly convergent and belongs to L_2. Hence we may integrate the series $\sum_{-\infty}^{\infty} g_\nu(u)\, e^{iux}$ termwise. Thus:

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(u)\, e^{iux}\,du \\ &= \frac{1}{\sqrt{2\pi}} \sum_{-\infty}^{\infty} \int_{a_\nu}^{b_\nu} \frac{e^{iux}\,du}{\sqrt{2\pi}} \\ &= \frac{1}{2\pi i x} \sum_{-\infty}^{\infty} (e^{ib_\nu x} - e^{ia_\nu x}). \end{aligned}$$

THE RANGE OF DE LA VALLÉE POUSSIN SUMMATION

By George Rutledge and R. D. Douglass

If the terms of the series

$$u_0 + u_1 + u_2 + u_3 + u_4 + \cdots, \tag{1}$$

and the terms of the series

$$u_0 + \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \cdots, \tag{2}$$

are related by means of the equations

$$u_0 + \sum_{j=1}^{n} \frac{\binom{2n}{n-j}}{\binom{2n}{n}} u_j = u_0 + \sum_{j=1}^{n} \lambda_j, \tag{3}$$

where $n = 1, 2, \cdots$, then the series (1) is summable by the method of de la Vallée Poussin,[1] or briefly, (VP) summable to the sum s when the series (2) converges to the sum s in the ordinary sense.

It has been previously shown[2] that the transformation (3) has the remarkably simple inverse

$$u_n = (-1)^{n-1} 2 \sum_{s=0}^{n-1} \binom{n-1}{s}\binom{n+s}{s}(-1)^s \lambda_{s+1}, \tag{4}$$

where $n = 1, 2, \cdots$. Thus if $\lambda_{s+1} = (-1)^s$ the series (1) and (2) become, respectively,

[1] See T. H. Gronwall, Summation of Series and Conformal Mapping, Annals of Math., ser. 2, v. 33 (1932), pp. 101–117.

In applying the de la Vallée Poussin transformation (3) we shall take $u_0 = 0$. Then the series $A + B + \cdots$ is transformed by (3) into the series

$$\left(\frac{A}{2}\right) + \left(\frac{A}{6} + \frac{B}{6}\right) + \cdots.$$

This is illustrated by the transformation of series (6) and (8) into series (7) and (9), respectively.

[2] Douglass, Stirling Expansions Derived by Means of Finite de la Vallée Poussin Summation, this Journal, v. 10 (1931), p. 135; see also Rutledge, The Inverse Matrix for de la Vallée Poussin Summation, this Journal, v. 11 (1932), p. 76.

$$2 - 8 + 38 - 192 + 1002 - 5336 + \cdots, \tag{6}$$

$$1 - 1 + 1 - 1 + 1 - 1 + \cdots, \tag{7}$$

while if $\lambda_{s+1} = (-1)^s \frac{1}{s+1}$ the series (1) and (2) become, respectively,

$$2 - \frac{10}{2} + \frac{50}{3} - \frac{258}{4} + \frac{1362}{5} - \frac{7306}{6} + \cdots, \tag{8}$$

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots. \tag{9}$$

The series (8) is summable by the transformation (3), while the series (6) is not. In a paper accepted for publication by the American Mathematical Monthly[3] the authors discuss the properties of two integral functions $Q(z)$, $M(z)$ which take positive integer values for $z = 1, 2, \cdots$. The general term of the series (6) is $(-1)^{n-1}Q(n)$, and the general term of the series (8) is $(-1)^{n-1}\frac{2M(n)}{n}$. It is shown that as n becomes infinite

$$\frac{n}{\sqrt{2}}\left\{\frac{2M(n)}{n}\right\} \sim Q(n) \sim \sqrt[4]{2}\,\frac{\cosh \alpha n}{\sqrt{\pi n}}, \tag{10}$$

where $\alpha = 2\sinh^{-1} 1$. Thus the ratio of the absolute value of the general term of the (VP) summable series (8) to the absolute value of the general term of the series (6) approaches zero as n becomes infinite. In this paper we propose to prove that $Q(n)$ defines the *range* of (VP) summation in the sense that any (VP) summable series necessarily has the ratio of the absolute value of its general term to $Q(n)$ approaching zero, but need not have the ratio of the absolute value of its general term to[4] $\frac{Q(n)}{n}$ approaching zero, as n becomes infinite. The theorem

[3] George Rutledge and R. D. Douglass, Integral Functions Associated with Certain Binomial Coefficient Sums.

[4] Moreover Douglass, loc. cit., p. 137, gives the example,

$$\sum_{n=1}^{\infty} (-1)^{n-1} \frac{\sinh \alpha n}{\sqrt{2}\, n} = 2 - \frac{12}{2} + \frac{70}{3} - \frac{408}{4} + \frac{2378}{5} - \frac{13860}{6} + \cdots,$$

of a (VP) summable series which has the ratio of the absolute value of its general term to $\frac{Q(n)}{\sqrt{n}}$ approaching the non-zero limit $\frac{\sqrt{\pi}}{2^{\frac{1}{2}}}$ as n becomes infinite.

which we shall prove is the following:

THEOREM. *If* $\sum_{j=1}^{n} \frac{\binom{2n}{n-j}}{\binom{2n}{n}} u_j$ *tends to a finite limit as* n *becomes infinite, then* $\frac{|u_n|}{Q(n)}$ *tends to the limit zero. In this sense the integral function* $Q(z)$ *defines the range of de la Vallée Poussin summation.*

It follows from (4) that for any series (1), whether (VP) summable or not, we have, for $n = 1, 2, \cdots$,

$$(11) \qquad \frac{|u_n|}{Q(n)} = \left| \frac{\sum_{s=0}^{n-1} \binom{n-1}{s}\binom{n+s}{s}(-1)^s \lambda_{s+1}}{\sum_{s=0}^{n-1} \binom{n-1}{s}\binom{n+s}{s}} \right| = |c_{n1}\lambda_1 + c_{n2}\lambda_2 + \cdots + c_{nn}\lambda_n|.$$

If we make the hypothesis that the series (1) is (VP) summable, then the series (2) is convergent and as a consequence λ_n tends to the limit zero as n becomes infinite.

Now the conditions (necessary and sufficient) under which

$$(12) \qquad c_{n1}\lambda_1 + c_{n2}\lambda_2 + \cdots + c_{nn}\lambda_n$$

approaches a finite limit as n becomes infinite in all cases in which λ_n approaches a finite limit as n becomes infinite are as follows:[5]

(i) that $\lim_{n\to\infty} c_{nk} = c_k$;

(ii) that $\lim_{n\to\infty} (c_{n1} + c_{n2} + \cdots + c_{nn}) = A$;

(iii) that $|c_{n1}| + |c_{n2}| + \cdots + |c_{nn}| < M$,

for every n, where M is a constant independent of n. Moreover when these conditions are satisfied

$$(13) \qquad \lim_{n\to\infty} (c_{n1}\lambda_1 + \cdots + c_{nn}\lambda_n) = A \cdot \lim_{n\to\infty} \lambda_n + \sum_{k=1}^{\infty} c_k(\lambda_k - \lim_{n\to\infty} \lambda_n).$$

[5] Kojima, Tôhoku Mathematical Journal, v. 12 (1917), pp. 291-326; J. Schur, Journal für Mathematik, v. 151 (1921), p. 82.

We note first that (iii) is satisfied by the coefficients c_{nk} in (11) because the sum of absolute values in question is equal to 1 for every n. Next by comparing coefficients of x^{n-1} in

$$(1 - x)^{n-1} \cdot (1 - x)^{-n-1} = (1 - x)^{-2}$$

we have

$$\sum_{s=0}^{n-1} (-1)^s \binom{n-1}{s} \binom{n+s}{s} = (-1)^{n-1} n, \tag{14}$$

and, since

$$\operatorname*{Lim}_{n\to\infty} 2 \frac{(-1)^{n-1} n}{Q(n)} = A = 0, \tag{15}$$

condition (ii) is satisfied. Finally we have

$$\operatorname*{Lim}_{n\to\infty} 2 \frac{(-1)^{k-1} \binom{n-1}{k-1} \binom{n+k-1}{k-1}}{Q(n)} = c_k = 0, \tag{16}$$

since the numerator is a polynomial of degree $2k - 2$ in n, and consequently condition (i) is satisfied. The conclusion of the theorem then follows immediately from (13). It may be noted that the convergence of the series (2) is sufficient but not necessary for the conclusion, since we obtain the zero limit from (13) on the hypothesis that $\operatorname*{Lim}_{n\to\infty} \lambda_n$ exists and is finite.

THE FOURIER TRANSFORM SOLUTION OF ORDINARY AND PARTIAL DIFFERENTIAL EQUATIONS[1]

By Norman Levinson

For a long time it has been known that the work of Heaviside in solving transient problems in electrical engineering and physics could be handled by the use of the Fourier transform. The Fourier transform point of view is a most fertile one, but seems to have been almost entirely neglected. Of the papers and text-books best known to the engineer only the treatment in Campbell and Foster, "Fourier Integral for Practical Applications," appears to be written with the point of view of using the Fourier transform to the fullest degree.

It is the purpose of this paper to introduce the Fourier integral and to use it to solve a number of problems of physical interest. The rather vague ideas of operators will be omitted entirely. The advantages of this method is that everything used will result from two theorems. Keeping these theorems always in view eliminates the necessity of using artifices and facilitates the setting up of important problems which have never been formulated by Heaviside's methods.

There are two possible methods of presentation. One is the logical and exact method of the pure mathematician which is not very intelligible to anyone who is not a mathematician. The second method makes no attempt to be entirely rigorous but gives a heuristic development of the subject which keeps the fundamental principles always in the foreground. It is the second method which we shall use here. The presentation will be subdivided in the following manner:

1. The Fourier Integral and Faltung Theorems.
2. The Ordinary Linear Differential Equation with Constant Coefficients or The Lumped Linear Network.
3. Partial Differential Equations of the Hyperbolic and Parabolic Type.
4. Non-Linear Systems.

[1] This is essentially a thesis accepted in June 1934 by the electrical engineering department of the Massachusetts Institute of Technology, as part of the requirement for the degree of Bachelor of Science and Master of Science.

1. The Fourier Integral and Faltung Theorems

It is well known that a function defined over a finite interval and satisfying certain general conditions can be expanded into a Fourier Series.[2] That is given $f(t)$ defined over $-l \leqq t \leqq l$ then

$$f(t) = \frac{1}{2b}\int_{-l}^{l} f(y)\,dy$$

$$+ \frac{1}{l}\sum_{n=1}^{\infty}\left(\sin\frac{n\pi t}{l}\int_{-l}^{l} f(y)\sin\frac{n\pi y}{l} + \cos\frac{n\pi t}{l}\int_{-l}^{l} f(y)\cos\frac{n\pi y}{l}\,dy\right)$$

$$= \frac{1}{\pi}\left[\frac{\pi}{2l}\int_{-l}^{l} f(y)\,dy + \sum_{n=1}^{\infty}\frac{\pi}{l}\int_{-l}^{l}\cos\frac{n\pi}{l}(t-y)f(y)\,dy\right]$$

$$= \frac{1}{\pi}\left[\frac{1}{2}\Delta u\int_{-l}^{l} f(y)\,dy + \sum_{n=1}^{\infty}\Delta u\int_{-l}^{l}\cos n\Delta u(t-y)f(y)\,dy\right],$$

$$\Delta u = \frac{\pi}{l}.$$

Now let $l \to \infty$ which is the same as $\Delta u \to 0$. Then the above equation *suggests*

$$f(t) = \frac{1}{\pi}\int_{0}^{\infty} du\int_{-\infty}^{\infty}\cos u(t-y)f(y)\,dy.$$

We assume that $\int_{-\infty}^{\infty}|f(y)|\,dy < \infty$. Note that we have not proved the formula but merely obtained it heuristically. A rigorous proof proceeds in much the same manner as in Fourier series.

Since

$$\int_{0}^{A}\cos x\,dx = \frac{1}{2}\int_{-A}^{A}\cos x\,dx,$$

we can write

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} du\int_{-\infty}^{\infty}\cos u(t-y)f(y)\,dy.$$

If we recall that $e^{ix} = \cos x + i\sin x$ and note that $\int_{-A}^{A}\operatorname{six} x\,dx = 0$,

[2] See "Diff. Equations for Electrical Engineers," by P. Franklin, p. 69.

we have

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} e^{iu(t-y)} f(y)\, dy$$

or

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iut}\, du \int_{-\infty}^{\infty} e^{iuy} f(y)\, dy.$$

If we now set

$$g(u) = \int_{-\infty}^{\infty} e^{-iuy} f(y)\, dy,$$

then we have

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(u)\, e^{iut}\, du.$$

This is the Fourier transform theorem.

If $\int_{-\infty}^{\infty} |f(t)|\, dt$ is not finite then $g(u)$ need not exist. In many cases of interest this situation will arise. For example a case of great interest is $f(t) = \sin \omega t$. Here no $g(u)$ is determined by the formula. In cases where transient effects are of interest, $f(t)$ is often a function for which $f(t) = 0$ for $t < 0$. Let c be a positive constant. Let us now work with $f_1(t) = e^{-ct} f(t)$. Applying the transform theorem to $f_1(t)$ we have

$$f_1(t) = e^{-ct} f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iut}\, du \int_{0}^{\infty} f(y)\, e^{-y(c+iu)}\, dy.$$

Let us set $p = iu + c$. Then

$$f(t)\, e^{-ct} = \frac{1}{2\pi} \int_{-i\infty+c}^{i\infty+c} e^{(p-c)t}\, du \int_{0}^{\infty} f(y)\, e^{-py}\, dy,$$

or

$$f(t) = \frac{1}{2\pi i} \int_{-i\infty+c}^{i\infty+c} e^{pt}\, dp \int_{0}^{\infty} f(y)\, e^{-py}\, dy.$$

If $\int_{0}^{\infty} |f(t)|\, dt$ is not bounded but $\int_{0}^{\infty} e^{-c_0 t} |f(t)|\, dt$ for some $c_0 > 0$ is bounded, then the above formula holds with $c > c_0$. Thus as a corollary

to the Fourier transform theorem we have if $f(t) = 0$ for $t < \infty$, and if $\int_0^\infty e^{-c_0 t} \mid f(t) \mid dt < \infty$ for some c_0, then if we define

$$g(p) = \int_0^\infty f(t) e^{-pt}\, dt$$

it follows that

$$f(t) = \frac{1}{2\pi i} \int_{-i\infty+c}^{i\infty+c} g(p) e^{pt}\, dp$$

with $c > c_0$. $g(p)$ is a function of the complex variable p and is known as the Laplace transform of $f(t)$. It is this form of the Fourier transform theorem that we will use in subsequent work. Whenever we use $g(p)$ we always assume that its transform $f(t) = 0$ for $t < 0$.

Let us now see what the transform of $f'(t) = \dfrac{df}{dt}$ is when $f(t) = 0$, $t < 0$.

Since only a continuous function can have a derivative it follows that $f(0) = 0$. Thus integrating by parts.

$$g(p) = \int_0^\infty f(t) e^{-pt}\, dt = \frac{1}{p} \int_0^\infty f'(t) e^{-pt}\, dt.$$

That is the transform of $f'(t)$ is $pg(p)$. If we denote that $g(p)$ is the transform of $f(t)$ by

$$f(t) \sim g(p),$$

then

$$f'(t) \sim pg(p).$$

Thus the operation of taking a derivative of a function is equivalent to multiplying its transform by p. This suggests using the transform on differential equations.

Let us try this on the simple equation

$$m \frac{d^2 s}{dt^2} + ks = f(t); \qquad (f(t) = 0, t < 0).$$

This is the dynamical equation of a mass, m, hanging in equilibrium from a spring up to $t = 0$, and with a force, $f(t)$, acting on the mass for $t > 0$. We set

$$g(p) = \int_0^\infty e^{-pt} s(t)\, dt \qquad \text{and} \qquad G(p) = \int_0^\infty e^{-pt} f(t)\, dt.$$

If we multiply the differential equation by e^{-pt} and integrate over $(0, \infty)$ with respect to t and use the fact that

$$s''(t) \sim p^2 g(p),$$

then we have

$$mp^2 g(p) + kg(p) = G(p).$$

Thus

$$g(p) = \frac{G(p)}{mp^2 + k}.$$

Or

$$s(t) \sim G(p) \frac{1}{mp^2 + k}.$$

We have expressed $s(t)$ as the transform of the product of two transforms. How can we express $s(t)$ in terms of $f(t) \sim G(p)$ and the transform of $\dfrac{1}{mp^2 + k}$? This problem is a fundamental one and can be formulated as follows.

Given $g_1(p) \sim f_1(t)$ and $g_2(p) \sim f_2(t)$, find $f_3(t) \sim g_1(p)g_2(p)$ in terms of $f_1(t)$ and $f_2(t)$. We have

$$f_3(t) = \frac{1}{2\pi i} \int_{-i\infty+c}^{i\infty+c} g_1(p)g_2(p)e^{pt}\, dp$$

$$= \frac{1}{2\pi i} \int_{-i\infty+c}^{i\infty+c} g_1(p)e^{pt}\, dp \int_0^\infty f_2(y)e^{-py}\, dy.$$

We now invert the order of integration in this last integral without questioning the validity of this step. We get

$$f_3(t) = \int_0^\infty f_2(y)\, dy \frac{1}{2\pi i} \int_{-i\infty+c}^{i\infty+c} g_1(p)e^{p(t-y)}\, dp.$$

But

$$\frac{1}{2\pi i} \int_{-i\infty+c}^{i\infty+c} g_1(p)e^{p(t-y)}\, dp = f_1(t - y).$$

Thus

$$f_3(t) = \int_0^\infty f_2(y)f_1(t - y)\, dy.$$

Since

$$f_1(t - y) = 0 \quad \text{for} \quad y > t,$$

$$f_3(t) = \int_0^t f_2(y)f_1(t - y)\, dy$$

If we set $x = t - y$ we have

$$f_3(t) = \int_0^t f_2(t - x) f_1(x)\, dx.$$

We have as our second theorem—
If[3] $f_1(t) \sim g_1(p)$, $f_2(t) \sim g_2(p)$ *and if* $f_3(t) \sim g_1(p)g_2(p)$, *then*

$$f_3(t) = \int_0^t f_1(t - \tau) f_2(\tau)\, d\tau.$$

This is the Faltung Theorem and is extremely important. It can be proved in a rigorous manner.

In the case of the differential equation we started to handle, if $\frac{1}{mp^2 + k} \sim f_1(t)$ then

$$s(t) = \int_0^t f(y) f_1(t - y)\, dy.$$

In this case $f_1(t) = \frac{1}{\sqrt{mk}} \sin \sqrt{\frac{k}{m}}\, t$ as can be verified from

$$\frac{1}{mp^2 + k} = \frac{1}{\sqrt{mk}} \int_0^\infty \sin \sqrt{\frac{k}{m}}\, t\, e^{-pt}\, dt.$$

Thus

$$s(t) = \frac{1}{\sqrt{mk}} \int_0^t \sin \sqrt{\frac{k}{m}}\, (t - y) f(y)\, dy.$$

2. The Ordinary Linear Differential Equation with Constant Coefficients or the Lumped Linear Network

Let us now apply the methods developed to a more general problem. We will handle a system of ordinary differential equations with constant coefficients. It will add to the interest if we use the terminology of some particular science. Let us therefore take an electrical problem. Note, however, that except for the terminology we might be solving a problem in mechanics or in some other field.

In the electrical case we get our equations by applying Kirchoff's rules to a network. If $q_k(t)$ is the charge that has passed through mesh k at the time t; if L_{km}, R_{km}, C_{km} are the mutual inductance; resistance, and capacitance of meshes k and m and $L_{kk}(t)$, R_{kk}, C_{kk} are the self inductances, resistance and capacitance of mesh k; and if $E_k(t)$ is the voltage applied to mesh k at the time t, then

[3] We recall that the symbol $\sim$ is read as "is the transform of."

$$\left(L_{11}\frac{d^2q_1}{dt^2} + R_{11}\frac{dq_1}{dt} + \frac{q_1}{C_{11}}\right) - \left(L_{12}\frac{d^2q_2}{dt^2} + R_{12}\frac{dq_2}{dt} + \frac{q_2}{C_{12}}\right)$$
$$- \cdots - \left(L_{1n}\frac{d^2q_n}{dt^2} + R_{1n}\frac{dq_n}{dt} + \frac{q_n}{C_{1n}}\right) = E_1(t),$$
$$- \left(L_{12}\frac{d^2q_1}{dt^2} + R_{12}\frac{dq_1}{dt} + \frac{q_1}{C_{12}}\right) + \left(L_{22}\frac{d^2q_2}{dt^2} + R_{22}\frac{dq_2}{dt} + \frac{q_2}{C_{22}}\right)$$
$$- \cdots - \left(L_{2n}\frac{d^2q_n}{dt^2} + R_{2n}\frac{dq_n}{dt} + \frac{q_n}{C_{2n}}\right) = E_2(t)$$
$$\vdots$$
$$- \left(L_{1n}\frac{d^2q_1}{dt^2} + R_{1n}\frac{dq_1}{dt} + \frac{q_1}{C_{1n}}\right) - \cdots + \left(L_{nn}\frac{d^2q_n}{dt^2} + R_{nn}\frac{dq_n}{dt} + \frac{q_n}{C_{nn}}\right)$$
$$= E_n(t).$$

Now since solutions of a linear system of equations are additive there is no loss in generality if we take $E_2(t) = E_3(t) = \cdots = E_n(t) = 0$. We now study the case where $E_1(t) = q_k(t) = 0$ for $t < 0$, and

$$\int_0^\infty |E_1(t)|\, e^{-c_0 t}\, dt < \infty \text{ for some } c_0 > 0.$$

Let us set

$$g_k(p) = \int_0^\infty q_k(t)e^{-pt}\, dt,$$
$$G_1(p) = \int_0^\infty E_1(t)e^{-pt}\, dt.$$

We have on multiplying the equations by e^{-pt} and integrating from 0 to ∞,

$$\left[p^2L_{11} + pR_{11} + \frac{1}{C_{11}}\right]g_1(p) - \left[p^2L_{12} + pR_{12} + \frac{1}{C_{12}}\right]g_2(p)$$
$$- \cdots - \left[p^2L_{1n} + pR_{1n} + \frac{1}{C_{1n}}\right]g_n(p) = G_1(p)$$
$$\cdots\cdots\cdots\cdots \qquad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots = 0$$
$$\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots$$
$$- \left[p^2L_{1n} + pR_{1n} + \frac{1}{C_{1n}}\right]g_1(p) - \cdots + \left[p^2L_{nn} + pR_{nn} + \frac{1}{C_{nn}}\right] = 0$$

Using determinants for solving we have

$$g_k(p) = \frac{\begin{vmatrix} \left(p^2L_{11} + pR_{11} + \frac{1}{C_{11}}\right) \cdots - (p^2L_{1k-1} + \cdots) & G_1(p) & -(p^2L_{1k+1} + \cdots) \cdots - \left(p^2L_{1n} + pR_{1n} + \frac{1}{C_{1n}}\right) \\ \cdot \quad \cdot & 0 & \cdot \quad \cdot \quad - \left(p^2L_{2n} + pR_{2n} + \frac{1}{C_{2n}}\right) \\ \cdot \quad \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ \left(p^2L_{n1} + pR_{n1} + \frac{1}{C_{n1}}\right) \cdots - (p^2L_{nk-1} + \cdots) & 0 & -(p^2L_{nk+1} + \cdots) \cdots + \left(p^2L_{nn} + pR_{nn} + \frac{1}{C_{nn}}\right) \end{vmatrix}}{\begin{vmatrix} \left(p^2L_{11} + pR_{11} + \frac{1}{C_{11}}\right) \cdots - (p^2L_{1k-1} + \cdots) - (p^2L_{1k} + \cdots) - (p^2L_{1k+1} + \cdots) \cdots - \left(p^2L_{1n} + pL_{1n} + \frac{1}{C_{1n}}\right) \\ \cdot \quad \cdot \quad \cdot \quad - \left(p^2L_{2n} + pR_{2n} + \frac{1}{C_{2n}}\right) \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ \cdot \quad \cdot \quad \cdot \quad \cdot \\ -\left(p^2L_{n1} + pR_{n1} + \frac{1}{C_{n1}}\right) \cdots - (p^2L_{nk-1} + \cdots) - (p^2L_{nk} + \cdots) - (p^2L_{1k+1} + \cdots) \cdots + \left(p^2L_{nn} + pR_{nn} + \frac{1}{C_{nn}}\right) \end{vmatrix}} = G_1(p)\,\frac{M_{1k}(p)}{D(p)}$$

Here $M_{1k}(p)$ is the minor of the 1st row and k^{th} column.
We can write $g_k(p) = G_1(p)F_{1k}(p)$ where

$$F_{1k}(p) = \frac{M_{1k}(p)}{D(p)}.$$

Using the transform theorem

$$q_k(t) = \frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} G_1(p)\, F_{1k}(p)\, e^{pt}\, dp.$$

If $F_{1k}(p) \sim A_{1k}(t)$ then by the Faltung Theorem

$$q_k(t) = \int_0^t dy\, E_1(y)\, A_{1k}(t-y) = \int_0^t E_1(t-y)\, A_{1k}(y)\, dy.$$

Or

$$\begin{aligned} i_k(t) = \frac{dq_k}{dt} &= E_1(t)A_{1k}(0) + \int_0^t E_1(y)A'_{1k}(t-y)\, dy \\ &= E_1(0)A_{1k}(t) + \int_0^t E'_1(t-y)A_{1k}(y)\, dy \\ &= E_1(t)A_{1k}(0) + \int_0^t E_1(t-y)A'_{1k}(y)\, dy \\ &= E_1(0)A_{1k}(t) + \int_0^t E'_1(y)A_{1k}(t-y)\, dy. \end{aligned}$$

$F_{1k}(p)$ may look more familiar if we note that

$$F_{1k}(p) = \frac{1}{pZ_{1k}(p)}$$

where Z_{1k} is an impedance function. That is $Z_{1k}(j\omega)$ is the impedance of the mesh k to a sinusoidal voltage of frequency ω applied to mesh 1. In other words, if we set $E_1(t) = e^{j\omega t}$ and solve the system of differential equations in the ordinary way for $I_{1k}(t)$, we find that

$$I_{1k}(t) = \frac{e^{j\omega t}}{Z_{1k}(j\omega)}.$$

If we set $E_1(t) = 1, t \geqq 0, E_1(t) = 0, t < 0$, then

$$i_k(t) = A_{1k}(t)E_1(0) + \int_0^t E'_1(t-y)A_{1k}(y)\, dy$$

becomes

$$i_k(t) = A_{1k}(t).$$

Thus $A_{1k}(t)$ has the physical significance in the electrical case of being the current that will flow in mesh k if a voltage $E_1(t) = 1$, $t > 0$ is applied to mesh 1. $A_{1k}(t)$ is the indicial admittance of Heaviside and the formulae for $i_k(t)$ in terms of $E_1(t)$ and $A_{1k}(t)$ are known as the superposition theorem in operational terminology.

We can express $A_{1k}(t)$ in a more explicit form. If we note that $M_{1k}(p)$ and $D(p)$ are polynomials in p and that

$$F_{1k}(p) = \frac{1}{pZ_{1k}(p)} = \frac{1}{p}\frac{pM_{1k}(p)}{D(p)}$$

then

$$A_{1k}(t) = \frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} \frac{e^{pt}}{pZ_{1k}(p)}\,dp$$

$$= \frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} \frac{c_0}{p}\frac{(p-c_1)\cdots(p-c_m)}{(p-a_1)\cdots(p-a_n)}\,e^{pt}\,dp.$$

We assume that $n \geqq m$. This will always be the case if $\int_0^\infty e^{-c_0 t}\,|\,A_{1k}(t)\,|\,dt$ is bounded for some $c_0 > 0$, because if it is bounded

$$F_{1k}(p) = \int_0^\infty e^{-pt}A_{1k}(t)\,dt,$$

must approach zero as $p \to \infty$, which means $F_{1k}(p)$ must contain less factors in its numerator than in its denominator.

If no $a_k = 0$, and if there are no equal a_k, it is a well-known theorem of algebra that

$$\frac{c_0(p-c_1)\cdots(p-c_m)}{p(p-a_1)\cdots(p-a_n)} = \frac{b_0}{p} + \frac{b_1}{p-a_1} + \cdots + \frac{b_n}{p-a_n}$$

where $b_0, b_1, \cdots, b_n$ are constants. This is known as a partial fraction expansion and is often used in evaluating integrals of algebraic functions. Note that if we multiply the expansion through by $p - a_l$ and set $p = a_l$ we have

$$b_l = \frac{c_0(a_l-c_1)\cdots(a_l-c_m)}{a_l(a_l-a_1)\cdots(a_l-a_{l-1})(a_l-a_{l+1})\cdots(a_l-a_n)}$$

$$= \frac{1}{a_l Z'_{1k}(a_l)} \qquad (l = 1, \cdots, n).$$

This can readily be seen by differentiating $pZ_{1k}(p)$ and noting that all except one term contain $(p - a_l)$, and therefore all except this term vanish for $p = a_l$. We also have

$$b_0 = \frac{c_0(-c_1)\cdots(-c_m)}{(-a_1)\cdots(-a_n)} = \frac{1}{Z_{1k}(0)}.$$

Therefore we have

$$\frac{1}{p(Z_{1k}(p))} = \frac{1}{pZ_{1k}(0)} + \frac{1}{a_1 Z'_1(a_1)}\frac{1}{p-a_1} + \cdots + \frac{1}{a_n Z'_{1k}(a_n)}\frac{1}{p-a_n}.$$

To get the transform of this expression we must find the transform of $\frac{1}{p}$. By integration $\int_0^\infty e^{-pt}\,dt = \frac{1}{p}$. Thus $\frac{1}{p} \sim f(t) = 1,\ t > 0$. Unless otherwise stated we are always concerned with functions of t which vanish for $t < 0$.

From this we now can readily compute

$$\begin{aligned}\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c}\frac{e^{pt}}{p+a+jb}\,dp &= \frac{1}{2\pi j}\int_{-j\infty+c+a}^{j\infty+c+a}\frac{e^{pt}-(a+jb)}{p}\,dp \\ &= e^{-(a+jb)t}\,\frac{1}{2\pi j}\int_{-j\infty+c_1}^{j\infty+c_1}\frac{e^{pt}}{p}\,dp \\ &= e^{-(a+jb)t}, \qquad t > 0,\end{aligned}$$

for any real a and b.

If we now use the partial fraction expansion for $\frac{1}{pZ_{1k}(0)}$ we have

$$A_{1k}(t) = \frac{1}{Z_{1k}(0)} + \frac{e^{a_1 t}}{a_1 Z'_{1k}(a_1)} + \cdots + \frac{e^{a_n t}}{a_n Z'_{1k}(a_n)}, \qquad t > 0.$$

This is the Heaviside expansion formula. In case of equal roots in $D(p) = 0$ we use the appropriate partial fraction expansion and obtain such terms as

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c}\frac{e^{pt}}{p^2}\,dp.$$

We have by use of the faltung theorem

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} e^{pt}\left(\frac{1}{p}\right)\left(\frac{1}{p}\right)dp = \int_0^t dy = t, \qquad (t > 0).$$

In this way all rational algebraic terms can be handled.

The discussion up to now has been limited to circuits which were dead for $t < 0$. This is really an inconsequential restriction that can readily be removed. To show this we handle the following example (Fig. 1) of a two mesh network with initial currents and charges. There is no distinction between this case and a more general network.

$$\left(L_{11}\frac{d^2}{dt^2} + R_{11}\frac{d}{dt} + \frac{1}{C_{11}}\right) q_1 - \left(L_{12}\frac{d^2}{dt^2} + R_{12}\frac{d}{dt} + \frac{1}{C_{12}}\right) q_2 = e_1(t),$$

$$-\left(L_{12}\frac{d^2}{dt^2} + R_{12}\frac{d}{dt} + \frac{1}{C_{12}}\right) q_1 + \left(L_{22}\frac{d^2}{dt^2} + R_{22}\frac{d}{dt} + \frac{1}{C_{22}}\right) q_2 = e_2(t).$$

For $t > 0$, $e_1(t)$ and $e_2(t)$ are given. We also know $q_1(0)$, $q_2(0)$, $q_1'(0)$ and $q_2'(0)$.

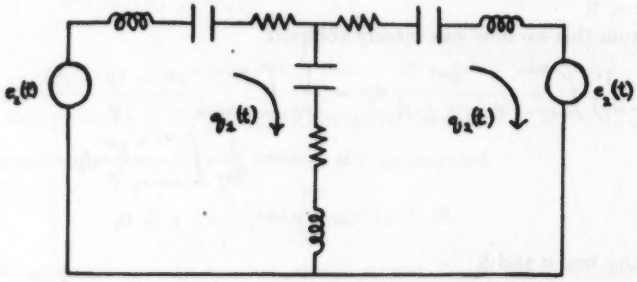


FIG. 1

Let

$$g_k(p) = \int_0^\infty e^{-pt} q_k(t)\, dt.$$

Integrating by parts we have

$$g_k(p) = \frac{q_k(0)}{p} + \frac{1}{p}\int_0^\infty e^{-pt}\frac{dq_k}{dt}\, dt.$$

Repeating this we get

$$g_k(p) = \frac{q_k(0)}{p} + \frac{q_k'(0)}{p^2} + \frac{1}{p^2}\int e^{-pt}\frac{d^2q_k}{dt^2}\, dt.$$

Or

$$\int_0^\infty e^{-pt} \frac{dq_k}{dt}\, dt = pg_k(p) - q_k(0)$$

$$\int_0^\infty e^{-pt} \frac{d^2q_k}{dt^2}\, dt = p^2g_k(p) - pq_k(0) - q_k'(0).$$

We can now multiply the mesh equations by e^{-pt}, integrate from 0 to ∞ in the usual way, and use the above results for $\int_0^\infty e^{-pt}q_k'(t)\, dt$ and $\int_0^\infty e^{-pt}q_k''(t)\, dt$ instead of the usual $pq_k(p)$ and $p^2g_k(p)$. We then solve the resulting equations for $g_1(p)$ and $g_2(p)$ and find $q_1(t)$ and $q_2(t)$ for $t > 0$ by taking the transforms of $g_1(p)$ and $g_2(p)$.

Another case of interest comes up when certain circuit elements are short circuited or in general when some change in elements is made. This case can be handled in the same way as that of a network with non-vanishing initial conditions. Another method of handling this case will be given in discussing the long line.

By use of the Fourier transform the ordinary linear differential equation with constant coefficients can be conveniently handled. This problem can be handled by other methods not involving the Fourier transform.[4] However, with the Campbell-Foster tables of Fourier transforms available, the Fourier transform method is most convenient. When we turn to partial differential equations the Fourier transform is not only convenient but indispensable.

3. Partial Differential Equations of the Hyperbolic and Parabolic Type

The partial differential equation of the hyperbolic type is exemplified by the long line equation while that of the parabolic type is exemplified by the heat flow equation.

The partial differential equation of a long line is[5]

$$\frac{\partial V}{\partial x} = -IR - L\frac{\partial I}{\partial t}$$

$$\frac{\partial I}{\partial x} = -GV - C\frac{\partial V}{\partial t}.$$

[4] See P. Franklin, Loc. cit., Chap. III.

[5] P. Franklin, Loc. cit., pp. 157, 158.

Eliminating I we have

$$\frac{\partial^2 V}{\partial x_2} = LC\frac{\partial^2 V}{\partial t^2} + (RC + LG)\frac{\partial V}{\partial t} + RGV.$$

The same equation holds for I. Using the Fourier transform theorem we set

$$V(x, t) = \frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} g(x, p)e^{pt}\,dp,$$

$$g(x, p) = \int_0^\infty V(x, t)p^{-pt}\,dt.$$

If we multiply the partial differential equation by e^{-pt} and integrate from 0 to ∞ with respect to t we have

$$\frac{\partial^2 g}{\partial x^2} = [LCp^2 + (RC + LG)\,p + RG]\,g.$$

Let us set

$$\alpha^2 = LCp^2 + (RC + LG)p + RG = (Lp + R)(Cp + G).$$

Then from the theory of ordinary differential equations

$$g(x, p) = A(p)e^{-\alpha x} + B(p)e^{\alpha x}$$

where $A(p)$ and $B(p)$ are any two functions of p.

Let $I(x, t) \sim h(x, p)$. Then in the same way

$$h(x, p) = C(p)e^{-\alpha x} + D(p)e^{\alpha x}.$$

If we now make use of $\dfrac{\partial V}{\partial x} = -L\dfrac{\partial I}{\partial t} - IR$, we get

$$-\alpha A(p)e^{-\alpha x} + \alpha B(p)e^{\alpha x} = -(R + Lp)[C(p)e^{-\alpha x} + D(p)e^{\alpha x}],$$

hence

$$C(p) = \frac{\alpha A(p)}{R + Lp}, \qquad D(p) = -\frac{\alpha B(p)}{R + Lp}.$$

Now let

$$Z_0 = \sqrt{\frac{R + Lp}{G + Cp}}.$$

Thus

$$C(p) = \frac{A(p)}{Z_0}, \qquad D(p) = \frac{-B(p)}{Z_0}$$

$$h(x, p) = \frac{A(p)}{Z_0} e^{-ax} - \frac{B(p)}{Z_0} e^{ax}.$$

Fig. 2 shows a generator working into a sending impedance, Z_s, which in turn works into a long line of length l. At the end of the line is the receiving impedance Z_r.

In this case the boundary conditions are

$$\begin{cases} V(0, t) = E(t) - V_s \\ V(l, t) = V_R \end{cases}$$

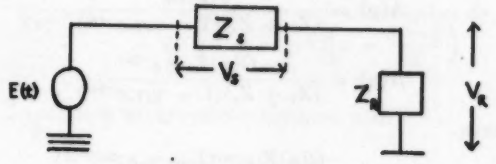


FIG. 2

To get these conditions in terms of the transform we multiply by e^{-pt} and integrate from 0 to ∞ with respect to t. Setting $F(t) \sim G(p)$ and using the results of Chapter II on lumped circuits to give

$$V_s \sim h(0, p)Z_s(p)$$

the boundary conditions are

$$g(0, p) = G(p) - h(0, p)Z_s(p),$$
$$g(l, p) = h(l, p)Z_r(p).$$

These conditions may be written as

$$A(p) + B(p) = G(p) - \left(\frac{A(p)}{Z_0} - \frac{B(p)}{Z_0}\right) Z_s$$

$$A(p)e^{-al} + B(p)e^{al} = \left(\frac{A(p)}{Z_0} e^{-al} - \frac{B(p)}{Z_0} e^{al}\right) Z_r.$$

Or

$$A(p)(Z_0 + Z_s) + B(p)[Z_0 - Z_s] = G(p)Z_0,$$

$$A(p)[Z_0 - Z_r]e^{-al} + B(p)[Z_0 + Z_r]e^{al} = 0;$$

$$A(p) = \frac{G(p)e^{al}(Z_0 + Z_r)Z_0}{(Z_0 + Z_s)(Z_0 + Z_r)e^{al} - (Z_0 - Z_s)(Z_0 - Z_r)e^{-al}},$$

$$B(p) = \frac{-G(p)e^{-al}(Z_0 - Z_r)Z_0}{(Z_0 + Z_s)(Z_0 + Z_r)e^{al} - (Z_0 - Z_s)(Z_0 - Z_r)e^{-al}}.$$

Let us set

$$\gamma_r = \frac{Z_r - Z_0}{Z_r + Z_0}, \qquad \gamma_s = \frac{Z_s - Z_0}{Z_s + Z_0}.$$

Then

$$A(p) = \frac{G(p)Z_0}{(Z_0 + Z_s)(1 - \gamma_s\gamma_r e^{-2al})},$$

$$B(p) = \frac{-G(p)Z_0\gamma_r e^{-2al}}{(Z_0 + Z_s)(1 - \gamma_s\gamma_r e^{-2al})}.$$

We now have

$$g(x, p) = \frac{G(p)Z_0e^{-ax}(1 - \gamma_r e^{-2a(l-x)})}{(Z_0 + Z_s)(1 - \gamma_r\gamma_s e^{-2al})},$$

$$h(x, p) = \frac{G(p)e^{-ax}(1 + \gamma_r e^{-2a(l-x)})}{(Z_0 + Z_s)(1 - \gamma_r\gamma_s e^{-2al})},$$

and we recall that $V(x, t) \sim g(x, p)$ and $I(x, t) \sim h(x, p)$.

For use in faltung formula we would have a satisfactory solution if we obtained

$$B(x, t) = \frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} \frac{e^{pt}Z_0(p)e^{-ax}(1 - \gamma_r e^{-2a(l-x)})}{p(Z_0 + Z_s)(1 - \gamma_r\gamma_s e^{-2al})}\,dp$$

for then

$$V(x, t) = \frac{d}{dt}\int_0^t B(x, \tau)E(t - \tau)\,d\tau.$$

Let us take the case where $\gamma_r = \gamma_s = 0$. This means that $Z_0 = Z_s = Z_r$ and is of interest in communication systems. Then we are concerned with

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} \frac{e^{pt}e^{-ax}}{2p}\,dp$$

for V and

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c}\frac{e^{pt}e^{-ax}}{2pZ_0(p)}\,dp$$

for I. Note that except for a constant factor the formulas are the same as for an infinite line with $Z_s = 0$.

First we will concern ourselves with

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c}\frac{e^{pt}e^{-x\sqrt{(p+a)(p+b)}}}{\sqrt{(p+a)(p+b)}}\,dp$$

Let $p = s - \frac{1}{2}(a+b)$. Then we have

$$\frac{e^{-\frac{1}{2}(a+b)t}}{2\pi j}\int_{-j\infty+c}^{j\infty+c}\frac{e^{-st}e^{-x\sqrt{\left(s+\frac{a-b}{2}\right)\left(s-\frac{a-b}{2}\right)}}}{\sqrt{\left(s+\frac{a-b}{2}\right)\left(s-\frac{a-b}{2}\right)}}\,ds.$$

If we set $\frac{a-b}{2} = n$ we are now concerned with

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c}\frac{e^{pt}e^{-x\sqrt{p^2-n^2}}}{\sqrt{p^2-n^2}}\,dp.$$

We now introduce integration in the complex plane to evaluate this function. We will show that it is a Bessel function. One who is not familiar with the Bessel function or with integration in the complex plane can avoid all work by taking this as a definition.

For $t < x$ we choose c sufficiently large so that $c > n$. Clearly,

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c}\frac{e^{pt}e^{-x\sqrt{p^2-n^2}}}{\sqrt{p^2-n^2}}\,dp = \frac{1}{2\pi j}\lim_{A\to\infty}\int_{-jA+c}^{jA+c}\frac{e^{pt}e^{-x\sqrt{p^2-n^2}}}{\sqrt{p^2-n^2}}\,dp.$$

Now we close the contour on the right and consider the closed contour shown in Fig. 3. This path encloses no singularities, therefore the integral around it is zero. That is, if $p = c + Ae^{j\theta}$

$$\int_{-jA+c}^{jA+c}\frac{e^{pt}e^{-x\sqrt{p^2-n^2}}}{\sqrt{p^2-n^2}}\,dp + \int_{\frac{\pi}{2}}^{-\frac{\pi}{2}}\frac{jd\theta e^{(Ae^{j\theta}+c)t-x(Ae^{j\theta}+c)\sqrt{1-\frac{n^2}{(c+Ae^{j\theta})^2}}}}{\sqrt{\left(1+\frac{ce^{-j\theta}}{A}\right)^2-\frac{n^2e^{-2j\theta}}{A^2}}} = 0$$

For fixed x, c, and t, if A is sufficiently large we have

$$\left| \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\theta \frac{e^{ct} e^{Ae^{j\theta}t} e^{-x(Ae^{j\theta}+c)} \sqrt{1-\frac{n^2}{(c+Ae^{j\theta})^2}}}{\sqrt{\left(1+\frac{ce^{-j\theta}}{A}\right)^2 - \frac{n^2 e^{-2j\theta}}{A^2}}} \right|$$

$$\leqq \left| 4 \int_0^{\frac{\pi}{2}} e^{ct} e^{A\cos\theta(t-x)}\, d\theta \right| \leqq 4e^{ct} \int_0^{\frac{\pi}{2}} e^{-A(x-t)\cos\theta}\, d\theta$$

Thus for $t < x$,

$$\left| \frac{1}{2\pi j} \int_{-j\infty+c}^{j\infty+c} \frac{e^{pt-x\sqrt{p^2-n^2}}}{\sqrt{p^2-n^2}}\, dp \right| \leqq \lim_{A\to\infty} 4e^{ct} \int_0^{\frac{\pi}{2}} e^{-A(x-t)\cos\theta}\, d\theta = 0.$$

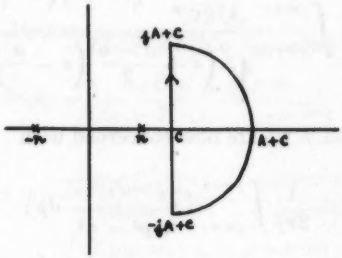


FIG. 3

For $t > x$ we rationalize by the substitution $p = \frac{n}{2}\left(w + \frac{1}{w}\right)$. We have

$$\frac{1}{2\pi j} \int_{-j\infty+c}^{j\infty+c} \frac{e^{pt-x\sqrt{p^2-n^2}}}{\sqrt{p_2-n^2}}\, dp = \frac{1}{2\pi j} \int_L \frac{dw}{w}\, e^{w\left(\frac{n(t-x)}{2}\right)+\frac{1}{w}\left(\frac{n(t+x)}{2}\right)}.$$

We get L, the path of integration from $w = \frac{p}{n} + \sqrt{\frac{p^2}{n^2}-1}$. We choose $c > n$ and deform the path of integration without crossing $w = 0$, and we get for L the line $-j\infty + c, j\infty + c$. We can again consider the path $(-jA + c, jA + c)$ for sufficiently large A, and we can close the path on the left half-plane. We can now close down to the pole at the origin. We get

$$\frac{1}{2\pi j} \int_{-j\infty+c}^{j\infty+c} \frac{dw}{w}\, e^{\frac{n}{2}\left[w(t-x)+\frac{1}{w}(t+x)\right]} = \frac{1}{2\pi j} \int_{|w|=k} \frac{dw}{w}\, e^{\frac{n}{2}\left[w(t-x)+\frac{1}{w}(t+x)\right]},$$

where k is any positive constant. Since $t > x$ we have if $\lambda = w\sqrt{\frac{t-x}{t+x}}$ that the integral becomes

$$\frac{1}{2\pi j}\int_{|\lambda|=k}\frac{d\lambda}{\lambda}\,e^{\frac{n}{2}\sqrt{t^2-x^2}\left(\lambda+\frac{1}{\lambda}\right)}.$$

For simplicity let us take the unit circle as the circle of integration. Then $\lambda = e^{i\theta}$ and we have

$$\frac{1}{2\pi}\int_0^{2\pi} e^{n\sqrt{t^2-x^2}\cos\theta}\,d\theta.$$

We can immediately show that this function is $I_0(n\sqrt{t^2-x^2})$ where $I_0(z) = J_0(iz)$ and J_0 is the zero order Bessel function. For integrating by parts we can readily show that

$$\frac{1}{2\pi}\int_0^{2\pi}\cos^{2n}\theta d\theta = \frac{2n-1}{2n}\,\frac{1}{2\pi}\int_0^{2\pi}\cos^{2n-2}\theta d\theta.$$

Using this and expanding $e^{z\cos\theta}$ into series form we have

$$\frac{1}{2\pi}\int_0^{2\pi} e^{z\cos\theta}\,d\theta = 1 + \frac{z^2}{2^2} + \frac{z^4}{(2!)^2(2)^4} + \cdots$$

and this is the familiar power series for $I_0(z)$. Thus

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c}\frac{e^{pt-x\sqrt{p^2-n^2}}}{\sqrt{p^2-n^2}} = \begin{cases} I_0(n\sqrt{t^2-x^2}), & t > x \\ 0 \quad\quad\quad\quad\quad , & t < x \end{cases}.$$

Or from the transformation made at the beginning

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c}\frac{e^{pt-x\sqrt{(p+a)(p+b)}}}{\sqrt{(p+a)(p+b)}}\,dp$$

$$= \begin{cases} e^{-\frac{1}{2}(a+b)t} I_0[\tfrac{1}{2}(a+b)\sqrt{t^2-x^2}], & t > x \\ 0, \quad t < x \end{cases}.$$

If we now use the faltung theorem on this, multiplying it by $1/p$, we have

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c}\frac{e^{pt-x\sqrt{(p+a)(p+b)}}}{p\sqrt{(p+a)(p+b)}}\,dp$$

$$= \begin{cases} \displaystyle\int_x^t e^{-\frac{1}{2}(a+b)\tau} I_0[\tfrac{1}{2}(a+b)\sqrt{\tau^2-x^2}]\,d\tau \\ 0, \quad t < x. \end{cases}$$

Now to get $I(x, t)$, we have

$$I(x,t) \sim \frac{e^{-\sqrt{LC}\times\sqrt{\left(p+\frac{R}{L}\right)\left(p+\frac{G}{C}\right)}}}{2p}\sqrt{\frac{Cp+G}{Lp+R}}$$

$$= \tfrac{1}{2}\sqrt{\frac{C}{L}}\frac{e^{-\sqrt{LC}\times\sqrt{\left(p+\frac{R}{L}\right)\left(p+\frac{G}{C}\right)}}}{\sqrt{\left(p+\frac{R}{L}\right)\left(p+\frac{G}{C}\right)}}\left(1+\frac{G}{C}\frac{1}{p}\right)$$

Note that if $t < x\sqrt{LC}$ then $I(x, t) = 0$ or in other words the current travels up to x in the time $t = x\sqrt{LC} = \frac{x}{v}$, where $v = \frac{1}{\sqrt{LC}}$ is the velocity of propagation of an electro-magnetic disturbance. Using $v = \frac{1}{\sqrt{LC}}$ and the transform of $I(x, t)$ as just written we have

$$I(x,t) = \begin{cases} \frac{1}{2}\sqrt{\frac{C}{L}}\, e^{-\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)t} I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)\sqrt{t^2-\frac{x^2}{v^2}}\right] \\ +\frac{1}{2}\frac{G}{\sqrt{LC}}\int_{\frac{x}{v}}^{t} e^{-\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)\tau} I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)\sqrt{\tau^2-\frac{x^2}{v^2}}\right]d\tau, \quad vt > x \\ 0, \qquad vt < x \end{cases}$$

$$V(x,t) \sim \frac{e^{-x\sqrt{(Lp+R)(Cp+G)}}}{2p}$$

$$= \frac{e^{-\frac{x}{v}\sqrt{\left(p+\frac{R}{L}\right)\left(p+\frac{G}{C}\right)}}}{2\sqrt{\left(p+\frac{R}{L}\right)\left(p+\frac{G}{C}\right)}}\,\frac{p^2+p\left(\frac{R}{L}+\frac{G}{C}\right)+\frac{RG}{LC}}{p\sqrt{\left(p+\frac{R}{L}\right)\left(p+\frac{G}{C}\right)}}$$

Now

$$\frac{1}{\sqrt{(p+a)(p+b)}} \sim \begin{cases} e^{-\frac{1}{2}(a+b)t} I_0(\tfrac{1}{2}(a+b)t), & t > 0 \\ 0, & t < 0, \end{cases}$$

by setting $x = v$ in the transform computed above. If we now use the faltung theorem we have

$$\frac{e^{-\frac{x}{v}\sqrt{\left(p+\frac{R}{L}\right)\left(p+\frac{G}{C}\right)}}}{\sqrt{\left(p+\frac{R}{L}\right)\left(p+\frac{G}{C}\right)}}\,\frac{1}{\sqrt{\left(p+\frac{R}{L}\right)\left(p+\frac{G}{C}\right)}}$$

$$\sim \int_{\frac{x}{v}}^{t} d\tau\, e^{-\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)\tau} I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)\sqrt{\tau^2-\frac{x^2}{v^2}}\right]$$

$$\times e^{-\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)(t-\tau)} I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)(t-\tau)\right]$$

$$= e^{-\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)t} \int_{\frac{x}{v}}^{t} d\tau I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)\sqrt{\tau^2-\frac{x^2}{v^2}}\right]$$

$$\times I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)(t-\tau)\right].$$

If we now go back to the transform of $V(x, t)$ we have remembering that $p \sim d/dt$ and that $1/p$ in the faltung theorem acts like $\int_0^t d\tau$,

$$V(x,t) = \frac{1}{4}\left(\frac{R}{L}+\frac{G}{C}\right) e^{-\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)t} \int_{\frac{x}{v}}^{t} d\tau I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)\sqrt{\tau^2-\frac{x^2}{v^2}}\right]$$

$$\times I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)(t-\tau)\right]$$

$$+ \frac{1}{2} e^{-\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)t} I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)\sqrt{t^2-\frac{x^2}{v^2}}\right]$$

$$+ \frac{RG}{2LC} \int_{\frac{x}{v}}^{t} d\xi e^{-\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)\xi} \int_{\frac{x}{v}}^{\xi} d\tau I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)\sqrt{\tau^2-\frac{x^2}{v^2}}\right]$$

$$\times I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)(\xi-\tau)\right] + \frac{1}{4}\left(\frac{R}{L}+\frac{G}{C}\right) e^{-\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)t}$$

$$\times \int_{\frac{x}{v}}^{t} d\tau I_0\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)\sqrt{\tau^2-\frac{x^2}{v^2}}\right] I_0'\left[\frac{1}{2}\left(\frac{R}{L}+\frac{G}{C}\right)(t-\tau)\right].$$

The order of integration can be changed in the double integral to $\int_{\frac{x}{v}}^{t} d\tau \int_{\frac{x}{v}}^{\tau} d\xi$. Note that if $G = 0$, as is usually the case in power transmission the formula for $I(x, t)$ involves no integration while that for $V(x, t)$ involves only a single integration.

Let us now consider the case of a line of length l with $Z_s = 0$ and $Z_r = \infty$ and with $G = 0$

Now

$$I(x,t) \sim \frac{e^{-ax}}{pZ_0}\frac{1+e^{-2a(l-x)}}{1+e^{-2al}}$$

or by setting

$$\frac{1}{1+e^{-2al}} = 1 - e^{-2al} + e^{-4al} + \cdots,$$

$$I(x,t) \sim \frac{1}{pZ_0}\left[e^{-ax} + e^{-a(2l-x)} - e^{-a(l+x)} + \cdots\right].$$

This expansion is legitimate since $|e^{-2al}| < 1$ along the path of integration. If we now use the formula

$$\frac{e^{-aA}}{pZ_0} \sim \sqrt{\frac{C}{L}}\, e^{-\frac{1}{2}\frac{R}{L}t}\, I_0\left[\frac{1}{2}\frac{R}{L}\sqrt{t^2 - \frac{A^2}{v^2}}\right], \qquad t > \frac{A}{v},$$

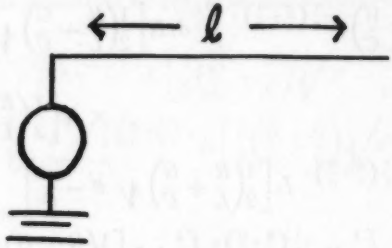


FIG. 4

then

$$I(x,t) = \sqrt{\frac{C}{L}}\, e^{-\frac{1}{2}\frac{R}{L}t}\left[I_0\left(\frac{1}{2}\frac{R}{L}\sqrt{t^2 - \frac{x^2}{v^2}}\right)_{t>\frac{x}{v}}\right.$$
$$+ I_0\left(\frac{1}{2}\frac{R}{L}\sqrt{t^2 - \frac{(2l-x)^2}{v^2}}\right)_{t>\frac{2l-x}{v}} - I_0\left(\frac{1}{2}\frac{R}{L}\sqrt{t^2 - \frac{(2l+x)^2}{v^2}}\right)_{t>\frac{2l+x}{v}}$$
$$+ I_0\left(\frac{1}{2}\frac{R}{L}\sqrt{t^2 - \frac{(4l-x)^2}{v^2}}\right)_{t>\frac{4l-x}{v}}$$
$$\left. - I_0\left(\frac{1}{2}\frac{R}{L}\sqrt{t^2 - \frac{(4l+x)^2}{v^2}}\right)_{t>\frac{4l+x}{v}} + \cdots\right]$$

$V(x, t)$ can be handled in a similar fashion. Notice that the terms behave like waves. The second term is a reflection from the open end. The third term is a reflection from the sending end, etc.

We now consider the heat flow equation. Let q be the heat flowing through a cross section of a rod, T the temperature, k the conductivity, c the heat capacity, and a the radiation constant. Then

$$q = -k\frac{\partial T}{\partial x},$$

$$\frac{\partial q}{\partial x} = -c\frac{\partial T}{\partial x} - aT,$$

$$k\frac{\partial^2 T}{\partial x^2} = c\frac{\partial T}{\partial t} + aT.$$

This last equation is exactly the same as the long line equation with no inductance. Thus the discussion of the long line applies here up to the point where the evaluation of integrals occur. Note that the results there involved the quantity $1/\sqrt{LC}$ which here is infinite. Thus we will start here by evaluating transforms. If we proceed exactly as in the electrical case we evaluate

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} \frac{e^{pt-x\sqrt{p+a}}}{\sqrt{p+a}}\,dp$$

$$= \lim_{A\to\infty} \frac{1}{2\pi j}\int_{-jA+c}^{jA+c} \frac{e^{pt-x\sqrt{p+a}}}{\sqrt{p+a}}\,dp$$

$$= \frac{e^{-at}}{2\pi j}\lim_{A\to\infty}\int_{-jA+c_1}^{jA+c_1} \frac{e^{pt-x\sqrt{p}}}{\sqrt{p}}\,dp.$$

For $t < 0$ we can close the path to the right since $e^{-x\sqrt{p}} \leqq e^{-x\sqrt{\frac{|p|}{2}}}$ for p to the right of the c_1 line and we have for $t < 0$ and $x > 0$,

$$\left|\frac{1}{2\pi}\int_{\frac{\pi}{2}}^{-\frac{\pi}{2}} \frac{e^{(Ae^{j\theta}+c_1)t-x\sqrt{Ae^{j\theta}+c_1}}}{\sqrt{Ae^{j\theta}+c_1}}\,Ae^{j\theta}\,d\theta\right| < \frac{1}{2\pi}\left|\int_{\frac{\pi}{2}}^{-\frac{\pi}{2}} \frac{Ae^{-x\sqrt{\frac{A}{2}}}}{\sqrt{A-c_1}}\,d\theta\right|$$

$$< \frac{1}{2}Ae^{-x\sqrt{\frac{A}{2}}}$$

And as $A \to \infty$ we see that the integral around the arc $(\pi/2, -\pi/2)$ approaches zero. Thus

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} \frac{e^{pt-x\sqrt{p+a}}}{\sqrt{p+a}}\,dp = 0, \qquad \text{for} \qquad t < 0.$$

For $t > 0$ we deform the path to the left. We make use of the fact that for $x > 0$

$$| e^{-x\sqrt{p}} | \leqq e^{-x\frac{\sqrt{|p|}}{2}}, \qquad 0 \leqq | \arg p | \leqq \frac{2}{3}\pi$$

where arg p is the angle of p. We also use the fact that

$$| e^{pt} | \leqq e^{-t\frac{|p|}{2}} \qquad \text{for} \qquad \pi \geqq | \arg p | \geqq \frac{2}{3}\pi .$$

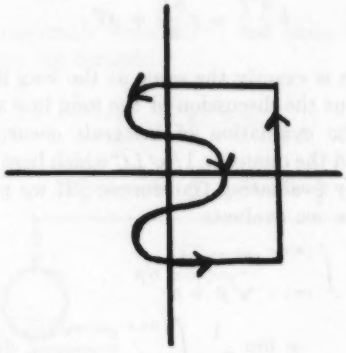


FIG. 5

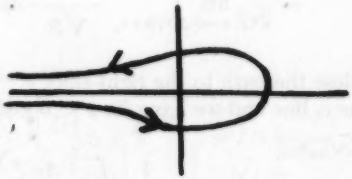


FIG. 6

If we now use the contour shown in Fig. 5, the facts just given allow us to discard the large arcs and we have that for $t > 0$, since the integral around the closed contour vanishes

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} \frac{e^{pt-x\sqrt{p}}}{\sqrt{p}}\,dp = \frac{1}{2\pi j}\int_{L} \frac{e^{pt-x\sqrt{p}}}{\sqrt{p}}\,dp$$

where L, Fig. 6, runs from $-\infty$ around the origin back to $-\infty$. Clearly

$$\frac{e^{-at}}{2\pi j}\int_L \frac{e^{pt-x\sqrt{p}}}{\sqrt{p}}\,dp = \frac{e^{-\frac{x^2}{4t}-at}}{2\pi j}\int_L \frac{e^{\left(\sqrt{pt}-\frac{x}{2\sqrt{t}}\right)^2}}{\sqrt{p}}\,dp$$

$$= \frac{e^{-\frac{x^2}{4t}-at}}{2\pi j\sqrt{t}}\int_L \frac{e^{\left(\sqrt{p}-\frac{x}{2\sqrt{t}}\right)^2}}{\sqrt{p}}\,dp.$$

If $p = y^2$ this becomes

$$\frac{e^{-\frac{x^2}{4t}-at}}{\pi j\sqrt{t}}\int_{-j\infty}^{j\infty} e^{\left(y-\frac{x}{2\sqrt{t}}\right)^2}\,dy.$$

If we set $jz = y - (x/2\sqrt{t})$ we have

$$\frac{e^{-\frac{x^2}{4t}-at}}{\pi\sqrt{t}}\int_{-\infty+\frac{jx}{2\sqrt{t}}}^{\infty+j\frac{x}{2\sqrt{t}}} e^{-z^2}\,dz.$$

Because of e^{-z^2} term this integral is the same along any parallel to the real axis. Thus we now have

$$\frac{e^{-\frac{x^2}{4t}-at}}{\pi\sqrt{t}}\int_{-\infty}^{\infty} e^{-z^2}\,dz.$$

But

$$\left[\int_{-\infty}^{\infty} e^{-z^2}\,dz\right]^2 = \int_{-\infty}^{\infty} e^{-x^2}\,dx\int_{-\infty}^{\infty} e^{-y^2}\,dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dx\,dy\,e^{-x^2-y^2} = \int_0^{2\pi} d\theta\int_0^{\infty} e^{-r^2}\,r\,dr = \pi.$$

Thus

$$\int_{-\infty}^{\infty} e^{-z^2}\,dz = \sqrt{\pi}.$$

Therefore for $x > 0$,

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} \frac{e^{pt-x\sqrt{p+a}}}{\sqrt{p+a}}\,dp = \begin{cases} \dfrac{e^{-\frac{x^2}{4t}-at}}{\sqrt{\pi t}}, & t > 0 \\ 0, & t < 0 \end{cases}.$$

In exactly the same way we can show

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} e^{pt-x\sqrt{p+a}}\,dp = \begin{cases} \dfrac{xe^{-\frac{x^2}{4t}-at}}{2\sqrt{\pi}\,t^{3/2}}, & t>0 \\ 0, & t<0 \end{cases}.$$

Other associated integrals can now be easily calculated. For example, to get

$$\frac{1}{2\pi j}\int_{-j\infty+c}^{j\infty+c} \frac{e^{pt-x\sqrt{p}}}{p}\,dp$$

we use the faltung theorem and we get

$$\int_0^t \frac{xe^{-\frac{x^2}{4\xi}}}{1\sqrt{\pi}\,\xi^{3/2}}\,d\xi.$$

If we set $\xi = x^2/4y^2$ in this integral we get

$$\frac{1}{\sqrt{\pi}}\int_{\infty}^{\frac{x}{2\sqrt{t}}} ye^{-y^2}\left(\frac{4y^2}{x^2}\right)\left(-\frac{2x^2}{4y^2}\right)dy$$

$$= \frac{2}{\sqrt{\pi}}\int_{\frac{x}{2\sqrt{t}}}^{\infty} e^{-y^2}\,dy = 1 - \frac{2}{\sqrt{\pi}}\int_0^{\frac{x}{2\sqrt{t}}} e^{-y^2}\,dy$$

$$= 1 - \operatorname{erf}\left(\frac{x}{2\sqrt{t}}\right).$$

As has already been mentioned the boundary problems here are set up just as in the long line and can be carried through in exactly the same way up to the point where the evaluation of integrals begins. At this point we proceed differently with the hyperbolic and parabolic type of equation as we have just seen.

Up to now we have discussed the handling of the transient problem for a distributed system, such as a long-line or heat-conducting rod, with the system disturbed only once at $t = 0$. The Fourier transform is very elastic and can readily be adapted to other cases.

First let us study a system which has been set in operation at $t = 0$, and in which some change is made at the time $t = t_0$. For example a long line might be set in operation at the time $t = 0$ and at $t = t_0$ some change is made in the terminal impedances. To find the current and voltage at any point and at any time we proceed as follows.

The change in the terminal impedances at $t = t_0$ is made by throwing some switches. In order to simplify our work we will handle the case

where one switch is closed at the receiving end of the line at the time $t = t_0$. All other cases are handled by similar means. (When a switch is opened care must be taken of the inductances in series with this switch by inserting some element to simulate the gradual discharge through an arc.)

The receiving end impedance must now be analysed. It has four terminals which interest us. There are the two terminals through which the line current enters and leaves the impedance and the pair of terminals which are short circuited by the closing of the switch at the time $t = t_0$. Let the impedance have n meshes and let mesh 1

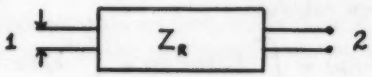


FIG. 7

involve the line circuit and mesh 2 the terminals to be short circuited (Fig. 7). Then we can write as the mesh equations

$$\left(L_{11}\frac{d}{dt^2}+R_{11}\frac{d}{dt}+\frac{1}{C_{11}}\right)q_1-\left(L_{12}\frac{d^2}{dt^2}+\cdots\right)q_2-\cdots-\left(L_{1n}\frac{d^2}{dt^2}+\cdots\right)q_n=V(l,t)$$

$$-\left(L_{12}\cdots\right)q_1+\left(L_{22}\cdots\right)q_2-\quad-\left(L_{2n}\right)q_n=?$$

$$\cdots\cdots\cdots

Using determinants to solve this lumped system as we have already done we can readily eliminate $h_2(p)$, $h_3(p)$, $\cdots$, $h_n(p)$ and we have left the equation

$$h(l, p) = \lambda(p)g(l, p) + u(p)G_0(p)$$

where $\lambda(p)$ and $u(p)$ are known.

As yet we have not obtained $E_0(t)$. It is very simple to do this. Solve the ordinary long line problem to which this problem degenerates if the mesh 2 is kept open for all time. If we do this we get $e_0(t)$ which is just the $E_0(t)$ we want for $t < t_0$, while for $t > t_0$ we have set $E_0(t) = 0$. Thus we can now calculate

$$G_0(p) = \int_0^\infty E_0(t)e^{pt}\,dp = \int_0^{t_0} e_0(t)e^{-pt}\,dt,$$

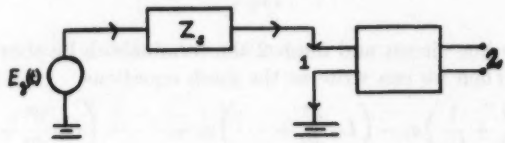


FIG. 8

directly or else we can use faltung theorem which gives, if $e_0(t) \sim g_0(p)$,

$$G_0(p) = \frac{1}{2\pi i}\int_{-i\infty+c}^{i\infty+c} \frac{1 - e^{-zt_0}}{z}\, g_0(p - z)\,dz.$$

Thus we have two ways to get $G_0(p)$ and we can use whichever is most convenient.

If we now look at the equation obtained for the receiving end

$$h(l, p) = \lambda(p)g(l, p) + u(p)G_0(p)$$

we see that it gives us a relation between $h(l, p)$ and $g(l, p)$ since all other terms are known.

We know from the solution already obtained in the ordinary case that

$$g(x, p) = A(p)e^{ax} + B(p)e^{-ax}$$

$$h(x, p) = \frac{A(p)e^{ax}}{Z_0} - \frac{B(p)}{Z_0}\,e^{ax}.$$

At the sending (Fig. 8) end we have

$$Gg(p) = g_0(p) + Z_s(p)h(0, p),$$

and we have already obtained the receiving end equation. If we now substitute for g and h we have

$$G_e(p) = A(p) + B(p) + \frac{Z_e(p)}{Z_0(p)}[A(p) - B(p)]$$

$$u(p)G_0(p) = \frac{A(p)e^{\alpha l} - B(p)e^{-\alpha l}}{Z_0} - \lambda(p)\{A(p)e^{\alpha l} + B(p)e^{-\alpha l}\}.$$

We can solve this system for $A(p)$ and $B(p)$ and thus get $g(x, p)$ and $h(x, p)$. To get the current and voltage, we find the transforms of $g(x, p)$ and $h(x, p)$.

In physical terminology what we have really done is added the currents which flow in each of the two systems shown below in Fig. 9.

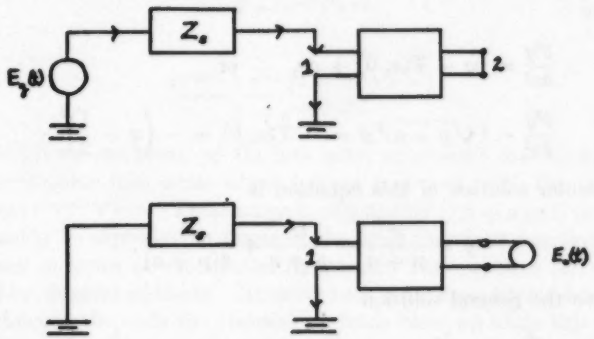


FIG. 9

Let us now turn to a different problem which also requires some discussion for solution by Fourier transforms. This is the problem that arises when the voltage and current along the line, or the temperature along a rod are not zero at the time $t = 0$. This problem can be easily handled. Let us take the case of a heat conducting rod of length l with an initial temperature distribution of

$$T(x, 0) = \left(x - \frac{l}{2}\right)^2.$$

Let us further take $T(0, t) = 0$, $T(l, t) = 0$. Let

$$g(x, p) = \int_0^\infty T(x, t)e^{-pt}\,dp = \frac{1}{p}T(x, 0) + \frac{1}{p}\int_0^\infty \frac{\partial T}{\partial t}e^{-pt}\,dp.$$

This integration by parts is most important. In fact once this step is taken the problem is solved. (In case of a long line two integrations by parts are necessary to bring in $\partial^2 V/\partial t^2$ term). We now multiply the equation

$$\frac{\partial^2 T}{\partial t^2} = \frac{\partial T}{\partial t} + aT$$

by e^{-pt} and integrate from 0 to ∞.

Having the first two coefficients unity is no restriction in generality. If we use

$$\int_0^\infty e^{-pt} \frac{\partial T}{\partial t}\, dt = pg(x, p) - T(x, 0)$$

we have

$$\frac{\partial^2 g}{\partial x^2} = pg - T(x, 0) + ag, \qquad \text{or}$$

$$\frac{\partial^2 g}{\partial x^2} - (\sqrt{p+a})^2 g = -T(x, 0) = -\left(x - \frac{l}{2}\right)^2.$$

A particular solution of this equation is

$$g = \frac{x^2}{p+a} - \frac{lx}{p+a} + \frac{l^2}{4(p+a)}.$$

We have the general solution

$$g(x, p) = \frac{x^2}{p+a} - \frac{lx}{p+a} + \frac{l^2}{4(p+a)}$$

$$+ \frac{2}{(p+a)^2}\left[A(p)e^{x\sqrt{p+a}} + B(p)e^{-x\sqrt{p+a}}\right]$$

where $A(p)$ and $B(p)$ are to be determined from boundary conditions. We have

$$g(0, p) = 0 = \frac{l^2}{4(p+a)} + \frac{2}{(p+a)^2} + A(p) + B(p),$$

$$g(l, p) = 0 = \frac{l^2}{4(p+a)} + \frac{2}{(p+a)^2} + A(p)e^{l\sqrt{p+a}} + B(p)e^{-l\sqrt{p+a}};$$

$$A(p) = \frac{-\left(\dfrac{l^2}{4(p+a)} + \dfrac{2}{(p+a)^2}\right)\left[e^{-l\sqrt{p+a}} - 1\right]}{e^{-l\sqrt{p+a}} - e^{l\sqrt{p+a}}},$$

$$B(p) = \frac{-\left(\frac{l^2}{4(p+a)} + \frac{2}{(p+a)^2}\right)[1 - e^{l\sqrt{p+a}}]}{e^{-l\sqrt{p+a}} - e^{l\sqrt{p+a}}}.$$

Thus

$$g(x,p) = \frac{(x-\frac{1}{2})^2}{p+a} + \frac{2}{(p+2)^2} + \left\{\frac{l^2}{4(p+a)} + \frac{2}{(p+a)^2}\right\} \times \frac{e^{-x\sqrt{p+a}}\{1 - e^{l\sqrt{p+a}}\} + e^{x\sqrt{p+a}}[e^{-l\sqrt{p+a}} - 1]}{e^{l\sqrt{p+a}} - e^{-l\sqrt{p+a}}}$$

If we consider

$$\frac{e^{-x\sqrt{p+a}}\{1 - e^{2(x-l)\sqrt{p+a}}\}}{1 - e^{-2l\sqrt{p+a}}}$$

and

$$\frac{e^{-(x+l)\sqrt{p+a}}\{1 - e^{2x\sqrt{p+a}}\}}{1 - e^{-2l\sqrt{p+a}}}$$

into which we can break up the last term, we observe that these terms are developable into series which behave very well, with the terms of the form $e^{-(x+l)\sqrt{p+a}}$. These terms multiplied by $1/p + a$ or $1/(p+a)^2$, can readily be expressed in terms of the error functions mentioned before, and in terms of exponential functions. This problem can also be solved by classical methods. However with more complicated boundary conditions at the ends the classical methods blow up while this attack does not. Note that in all problems of this type the most important thing is an integration by parts

$$\int_0^\infty T(x,t)e^{-pt}\,dt = \frac{1}{p}T(x,0) + \frac{1}{p}\int_0^\infty \frac{\partial T}{\partial t}e^{-pt}\,dt.$$

In connection with the long line the type of boundary conditions just discussed comes up in such cases as a charged line, etc. In the case of the vibrating chord this is a most natural form for the boundary conditions and in case there is damping or anything but the simplest end conditions the use of the Fourier transform will be necessary.

An important consideration is that in the case of a partial differential equation of the hyperbolic or parabolic type where the coefficients are functions of x but not of t the usual procedure of writing the transform equation gives an ordinary differential equation in x with non-constant coefficients which can sometimes be handled.

4. Non-Linear Systems

Up to now we have discussed linear systems only. By means of the Fourier transform it is possible to handle very special types of non-linear problems. For example let us consider the differential equation

$$a_0 \frac{d^n y}{dx^n} + a_1 \frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_{n-1} \frac{dy}{dt} + a_n y + \psi\left(\frac{dy}{dt}\right) = f(t).$$

Let us set

$$\phi(p) = \int_0^\infty y(t) e^{-pt}\, dt$$

$$\theta(p) = \int_0^\infty \psi\left(\frac{dy}{dt}\right) e^{-pt}\, dt$$

$$g(p) = \int_0^\infty f(t) e^{-pt}\, dt.$$

This implies that we are solving the problem for which $y(t) = 0$, $t < 0$ and $f(t) = 0$, for $t < 0$. The equation in terms of the transform becomes

$$F(p)\phi(p) + \theta(p) = g(p)$$

where

$$F(p) = a_0 p^n + a_1 p^{n-1} + \cdots + a_{n-1} p + a_n.$$

We have

$$\phi(p) = \frac{1}{F(p)} \{g(p) - \theta(p)\}$$

or

$$g(t) = \frac{1}{2\pi j} \int_{-j\infty+c}^{j\infty+c} \frac{g(p)}{F(p)} e^{pt}\, dp - \frac{1}{2\pi j} \int_{-j\infty+c}^{j\infty+c} \frac{\theta(p)}{F(p)} e^{pt}\, dp.$$

Note that $\dfrac{1}{2\pi j} \displaystyle\int_{-j\infty+c}^{j\infty+c} dp\, e^{pt} \frac{g(p)}{F(p)}$ is the solution to the problem if the $\psi\left\{\dfrac{dy}{dt}\right\}$ term is omitted. Let us call this term $y_0(t)$. Then

$$y(t) = y_0(t) - \frac{1}{2\pi j} \int_{-j\infty+c}^{j\infty+c} e^{pt} \frac{\theta(p)}{F(p)}\, dp.$$

Let $B(t) \sim 1/F(p)$. Then by faltung theorem

$$y(t) = y_0(t) - \int_0^t B(t-\tau)\psi\left\{\frac{dy(\tau)}{d\tau}\right\} d\tau.$$

We now differentiate and we have

$$\frac{dy}{dt} = \frac{dy_0}{dt} - \int_0^t B'(t-\tau)\psi\left\{\frac{dy(\tau)}{d\tau}\right\} d\tau - B(0)\psi\left\{\frac{dy(t)}{dt}\right\}.$$

This is an integral equation in dy/dt.

A method of procedure is the use of successive approximations to get dy/dt. There is always the question of convergence of the solution under successive approximations. This will depend upon the form of $\psi(dy/dt)$. Usually very rough methods will show whether or not the solution will converge.

Note that if instead of a term $\psi(dy/dt)$ we had a term $\psi(y)$ we would proceed just as above except that we would not differentiate at the very end but would leave the final equation in terms of $y(t)$ and $y_0(t)$.

Naturally the entire above procedure would work with a system of equations, one equation of which had a non-linear term, just as well as with a single equation.

A NOTE ON THE HEAWOOD COLOR FORMULA

By I. N. Kagno*

P. J. Heawood[1] derived a formula for an upper limit P_k to the number of mutually touching regions on a surface of characteristic K. This number is also an upper limit to the values of N for which a complete N-point can be mapped on the surface.[2] Philip Franklin[3] has pointed out that for the case of non-orientable surfaces with $K = 0$, the limit given by Heawood's formula cannot be attained, and asserts that

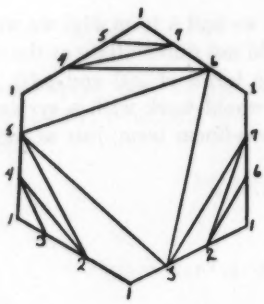


Fig. 1

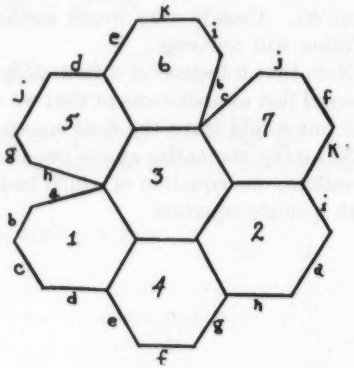


Fig. 2

possibly the same may happen for other values of K. Examples where the Heawood limit is attained are known for all orientable surfaces up to $K = -10$, and for the non-orientable surface with $K = 1$.[4] In this note we give examples for the non-orientable surfaces with $K = -1$, -2, and -4. As pointed out by Franklin[3] the color problem is then solved for these surfaces.

* Columbia University, New York City.

[1] P. J. Heawood—Quarterly Journal of Mathematics, Vol. 24 (1890), p. 332.

[2] For a statement of the duality between the problem of mapping complete N-points, and the problem of mapping N mutually touching regions on a surface, see Heffter—Math. Annalen, Vol. 38 (1891), p. 447.

[3] "A Six Color Problem"—Journal of Math. and Physics, Vol. 13 (1934), p. 363.

[4] Heffter, loc. cit.

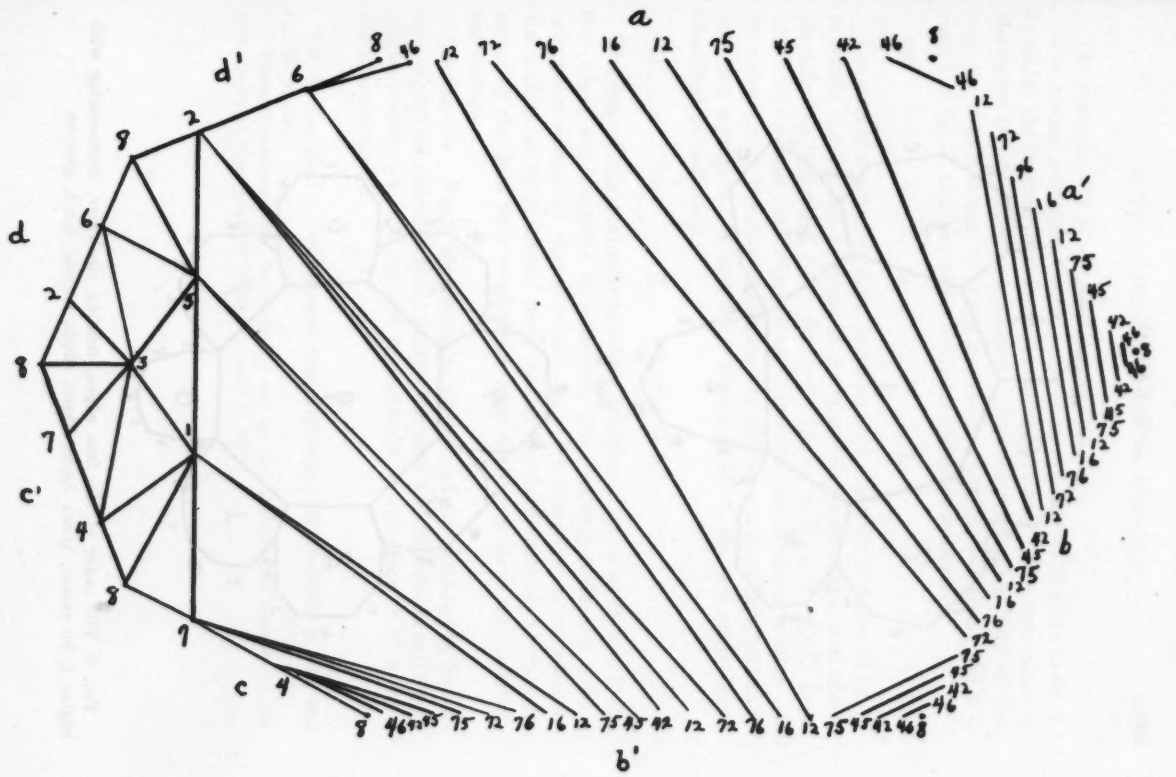


FIG. 3

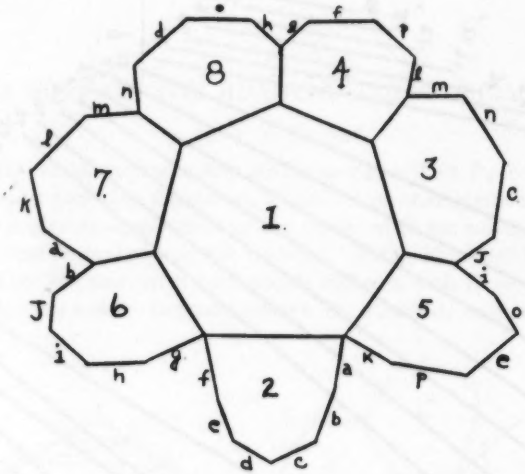


FIG. 4

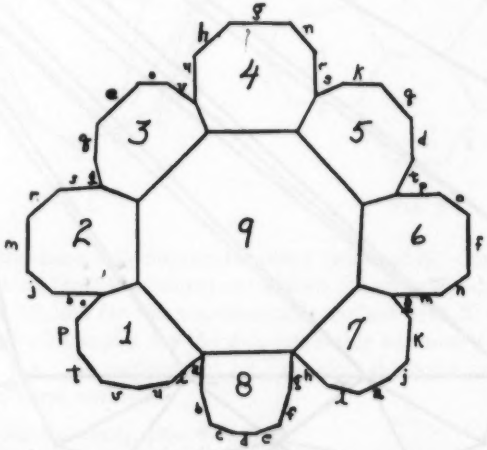


FIG. 5. The order of the free edges, reading clockwise, commencing with region 1, is: *iuvtpc, bjmrs, lqeov, uhgnr, skqdl, pofnml, kjaih, gfedcba.*

The examples (1) $K = -1$, $P_{-1} = 7$.

The map of a complete seven point on such a surface is given in fig. 1, in which the surface is given in its canonical form.[5] The free edges of the figure are identified as indicated. By constructing the dual[2] of this map we obtain a map of seven mutually touching hexagons. This is shown in fig. 2, the free edges being identified as indicated.

(2) $$K = -2, P_{-2} = 8.$$

The map of a complete eight point on such a surface is shown in fig. 3, in which the surface is given in its canonical form. The free edges, identified as indicated, have the order $(a, a', b, b', c, c', d, d')$ from which we see that the surface is non-orientable and of characteristic -2. By constructing the dual of this map we obtain a map of eight mutually touching septagons. This map is shown in fig. 4, the free edges being identified as indicated.

(3) $$K = -4, P_{-4} = 9.$$

The map of nine mutually touching octagons is given in fig. 5. If the free edges are identified as indicated, the resulting surface is a non-orientable surface of characteristic -4. We can affect a reduction of the surface to its canonical form by the standard methods.[6] By constructing the dual map we obtain a map of the complete nine point on the surface.

Conclusions. From the examples given, and the remarks in Franklin's paper[3] we conclude that; every map on a non-orientable surface of characteristic -1, -2, or -4 respectively, is colorable in 7, 8, or 9 colors respectively.

[5] B. v. Kerekjarto, "Vorlesungen uber Topologie," Berlin 1923, p. 155, and fig. 33, p. 156.

[6] These operations have been carried out by the author, but because of the complexity of the resulting figure we omit it here.

APPLICATIONS OF TAUBERIAN THEOREMS TO THE STUDY OF ENTIRE FUNCTIONS

By Bernard Friedman

In this paper I wish to give some applications of Professor Wiener's Tauberian Theorems[1] to the study of entire functions. Besides using Professor Wiener's original theorems I will use an extension due to Bochner, which is more convenient for future use.

Bochner[2] calls $\phi_p (p = 1, 2, \cdots)$ the class of all functions $\varphi(t)$ with the following properties in $0 < t < \infty$.

I. $\phi(t)$ is positive, continuous and 1 for $0 \leqq t \leqq 1$.

II. $\lim_{t\to\infty} \frac{\phi(e^t)}{\phi(t)} = 1$ uniformly in any interval $0 < c_0 \leqq c \leqq c_1$.

III. there is a constant C such that

$$\sum_{n=-\infty}^{\infty} \frac{\phi(e^n)}{1 + |x - n|^{p+1}} \leqq c\phi(e^x), \qquad -\infty < x < \infty.$$

If on putting $\psi(x) = \phi(e^x)$ and if $\max \psi(\xi) = O[\psi(x)]$ for $\frac{1}{2}x \leqq \xi \leqq x$ and if there are real numbers $\rho, \sigma, (\rho > 0)$ for which

$$\psi(x) = O(x^\rho), \qquad \psi^{-1}(x) = O(x^{-\sigma})$$

then III is fulfilled if $p \geqq \max(1 + \rho, \rho - \sigma)$.

All functions $\phi(x) = (\log x)^{\sigma_1}(\log_2 x)^{\sigma_2} \cdots (\log_n x)^{\sigma_n}$ belong to ϕ_p where $p = 1 + \sigma_1 + \epsilon$.

Bochner also calls $K_q (q = 0, 1, 2 \cdots)$ the class of all continuous functions which are $O(x^{-q})$ for $|x| \to \infty$.

Then the following lemma is a trivial extension of his Theorem I:

Lemma I. *Let there be given a real number $\gamma > 0$; a non-negative non-decreasing function $A(t)$; a function of class ϕ_p and a non-negative continuous function $W(t)$ for which*

$$t^\gamma W(t) = O[(\log t)^{-p-4}]$$

for both small and large t and

$$\int_0^\infty W(t)t^{\gamma-1+i\alpha}\,dt \neq 0, \qquad (-\infty < \alpha < \infty).$$

[1] Annals of Mathematics (2) 33, 1932, pp. 1–100.

[2] Journal of London Mathematical Society 9, 1934, pp. 141–149.

Suppose further that

$$\lim_{\lambda\to\infty} \lambda^{-\gamma}\,[\phi(\lambda)]^{-1} \int_0^\infty W\left(\frac{\mu}{\lambda}\right) dA(\mu) = A \int_0^\infty W(t) t^{\gamma-1}\, dt$$

and that the left hand function is bounded in $0 < \lambda < \infty$. *Then*

$$\lim_{\lambda\to\infty} \lambda^{-\gamma}\,[\phi(\lambda)]^{-1} \int_0^\infty F\left(\frac{\mu}{\lambda}\right) dA(\mu) = A \int_0^\infty F(t) t^{\gamma-1}\, dt$$

for any continuous $F(t)$ *which is for both small and large* t

$$O\,[t^{-\gamma}\,(\log t)^{-p-2}]$$

The following lemma is Bochner's Theorem III:

LEMMA II. *Under the above hypotheses*

$$\gamma A(\lambda) \sim A\lambda^{\gamma}\phi(\lambda) \text{ as } \lambda \to \infty.$$

The converse of this is

LEMMA III. *If*

$$\gamma A(\lambda) \sim A\lambda^{\gamma}\phi(\lambda) \text{ as } \lambda \to \infty$$

then

$$\lim_{\lambda\to\infty} \lambda^{-\gamma}\,[\phi(\lambda)]^{-1} \int_0^\infty F\left(\frac{\mu}{\lambda}\right) dA(\mu) = A \int_0^\infty F(t) t^{\gamma-1}\, dt$$

for any continuous function $F(t)$ *which is for both small and large* t

$$O\,[t^{-\gamma}\,(\log t)^{-p-2}]$$

Proof:

We can take $\gamma = 1$ and then get our lemma just as Bochner obtains his Theorem III from Theorem II. The hypothesis can be written

$$\lim_{\lambda\to\infty} \frac{1}{\lambda\phi(\lambda)} \int_0^\infty G\left(\frac{\mu}{\lambda}\right) dA(\mu) = A \int_0^\infty G(t)\, dt$$

where $G(t) = 1,\ 0 \leqq t \leqq 1;\ = 0,\ t > 1$. Let $G_2(t) = 1, 0 \leqq t \leqq 1 - \epsilon;\ = 0,\ t > 1 - \epsilon$.

Let $G_1(t) = 1,\ 0 \leqq t \leqq 1 - \epsilon;\ = 0,\ t > 1$; linear, $1 - \epsilon \leqq t \leqq 1$.

Now

$$\lim_{\lambda\to\infty}\frac{1}{\lambda\phi(\lambda)}\int_0^\infty G_2\left(\frac{\mu}{\lambda}\right)dA(\mu)=\lim_{\lambda\to\infty}\frac{1}{\lambda\phi(\lambda)}A\left[\lambda(1-\epsilon)\right]$$

$$=\lim_{\lambda\to\infty}\frac{1}{\lambda(1+\epsilon)\phi[\lambda(1+\epsilon)]}A(\lambda)=\frac{1}{1+\epsilon}\lim_{\lambda\to\infty}\frac{1}{\lambda\phi(\lambda)}A(\lambda)$$

$$=\frac{1}{1+\epsilon}A\int_0^\infty G(t)\,dt,\ \text{if}\ \frac{\phi[\lambda(1+\epsilon)]}{\phi(\lambda)}\to 1\ \text{as}\ \lambda\to\infty.$$

This will always be true by our hypothesis on ϕ. But $G_2 \leqq G_1 \leqq G$ and $G_2 \to G$ so

$$\lim_{\lambda\to\infty}\frac{1}{\lambda\phi(\lambda)}\int_0^\infty G_2\left(\frac{\mu}{\lambda}\right)dA(\mu)\leqq\overline{\lim_{\lambda\to\infty}}\frac{1}{\lambda\phi(\lambda)}\int_0^\infty G_1\left(\frac{\mu}{\lambda}\right)dA(\mu)$$

$$\leqq\lim_{\lambda\to\infty}\frac{1}{\lambda\phi(\lambda)}\int_0^\infty G\left(\frac{\mu}{\lambda}\right)dA(\mu)\ \therefore\ \lim_{\lambda\to\infty}\frac{1}{\lambda\phi(\lambda)}\int_0^\infty G_1\left(\frac{\mu}{\lambda}\right)dA(\mu)$$

$$=A\int_0^\infty G_1(t)\,dt$$

where $G_1(t)$ is continuous. Therefore the theorem follows by an application of Lemma I.

Let

$$E(u,0)=1+u.\quad E(u,m)=(1+u)e^{-u+\frac{u^2}{2}-\cdots(-1)^m\frac{u^m}{m}}\quad(m=1,2,\cdots)$$

Let $f(z)=\prod_{\nu=1}^{\infty}E\left(\frac{z}{a_\nu},m\right)$ where the a_ν are positive real numbers such that

$$0<a_1\leqq a_2\leqq\cdots\ \text{and}\ \sum_{\nu=1}^{\infty}a_\nu^{-m-1}<\infty.$$

Then we obtain the following theorem:[3]

THEOREM I. *Let $n(r)$ be the number of a_n not exceeding r. Let λ, ρ, θ, γ be fixed numbers such that $\lambda>0$, $0<\rho<1$, $|\theta|<\pi$, $\gamma=m+\rho$. Then the statements*

i. $n(x)\sim\lambda x^\gamma\phi(x)$,

ii. $\log f(x)\sim(-1)^m\pi\lambda\ \operatorname{cosec}\pi\rho\ x^\gamma\phi(x)$,

iii. $\log f(xe^{i\theta})\sim(-1)^m\pi\lambda\ \operatorname{cosec}\pi\rho\ e^{i\gamma\theta}x^\gamma\phi(x)$

are all equivalent.

[3] Paley and Wiener, Fourier Transforms in the Complex Domain, Theorem XXVI.

Proof:

i. can be written ($G(t)$ as in Lemma III)

$$\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\int_0^\infty G\left(\frac{t}{x}\right)dn(t)=\gamma\lambda\int_0^\infty G(t)t^{\gamma-1}\,dt$$

ii. can be written

$$\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\log f(x)=\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\int_0^\infty\left[\log\left(1+\frac{x}{t}\right)-\frac{x}{t}+\frac{x^2}{2t^2}-\cdots\right.$$

$$\left.(-1)^m\frac{x^m}{mt^m}\right]dn(t)=\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\int_0^\infty K_1\left(\frac{t}{x}\right)dn(t)=\gamma\lambda\int_0^\infty K_1(t)\,t^{\gamma-1}\,dt$$

Now

$$\int_0^\infty K_1(t)t^{\gamma-1+i\alpha}\,dt=\frac{K_1(t)t^{\gamma+i\alpha}}{\gamma+i\alpha}\Big|_0^\infty-\frac{1}{\gamma+i\alpha}\int_0^\infty K_1'(t)t^{\gamma+i\alpha}\,dt$$

$$=\frac{(-1)^m}{\gamma+i\alpha}\int_0^\infty\frac{t^{i\alpha+\rho-1}}{1+t}\,dt=\frac{(-1)^m}{\gamma+i\alpha}\pi\operatorname{cosec}\pi(i\alpha+\rho)\neq 0,\ (-\infty<\alpha<\infty).$$

In iii by separating into real and imaginary parts we have

$$\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\log|f(xe^{i\theta})|=\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\int_0^\infty\left[\frac{1}{2}\log\left|1+\frac{xe^{i\theta}}{t}\right|^2-\frac{x\cos\theta}{t}\right.$$

$$\left.+\frac{x^2\cos 2\theta}{2t^2}\cdots(-1)^m\frac{x^m\cos m\theta}{mt^m}\right]dn(t)$$

$$=\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\int_0^\infty K_2\left(\frac{t}{x}\right)dn(t)=\gamma\lambda\int_0^\infty K_2(t)\,t^{\gamma-1}\,dt.$$

Now

$$\int_0^\infty K_2(t)\,t^{\gamma-1+i\alpha}\,dt=\frac{(-1)^m\pi}{\gamma+i\alpha}\operatorname{cosec}\pi(i\alpha+\rho)\cos\theta(i\alpha+\gamma)\neq 0,$$

$$(-\infty<\alpha<\infty)$$

Also

$$\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\,\mathrm{I}\,[\log f(xe^{i\theta})]=\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\int_0^\infty\tan^{-1}\frac{x\sin\theta}{t+x\cos\theta}-\frac{x\sin\theta}{t}+\cdots$$

$$\left.+(-1)^m\frac{x^m\sin m\theta}{mt^m}\right]dn(t)$$

$$=\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\int_0^\infty K_3\left(\frac{t}{x}\right)dn(t)=\gamma\lambda\int_0^\infty K_3(t)\,t^{\gamma-1}\,dt.$$

Now

$$\int_0^\infty K_3(t)\, t^{i\alpha+\gamma-1}\, dt = \frac{-1}{\gamma + i\alpha} \int_0^\infty K_3'(t)\, t^{\gamma+i\alpha}\, dt$$

$$= \frac{(-1)^m \pi}{\gamma + i\alpha} \operatorname{cosec} \pi(i\alpha + \rho) \sin \theta(i\alpha + \gamma) \neq 0, \quad (-\infty < \alpha < \infty).$$

Now since $n(t)$ is a non-negative non-decreasing function and since $(-1)^m K_i(\lambda)$ after a fixed λ is a non-negative continuous function which satisfies the assumptions of Lemma I, our theorem follows by the application of Lemmas I, II, III.

This theorem enables us to determine approximately the points at which $f(z)$ is real. For $f(z)$ to be real the imaginary part of the logarithm must vanish or be a multiple of π. Hence for all x such that

$$\lambda \pi x^\gamma \phi(x) \operatorname{cosec} \pi\rho \sin \theta\gamma \sim \nu\pi,$$

$f(xe^{i\theta})$ is real. Let $1/\lambda \operatorname{cosec} \pi\rho = \kappa$ then $x^\gamma \phi(x) \sin \theta\gamma \sim \nu\kappa$ or on any sufficiently large circle $x =$ constant, there will be $2(m + 1)$ real points.

If $\phi(x) \equiv 1$ then along any line $\theta =$ constant the points of $f(z)$ which are real are at a distance of $(\kappa/\sin \theta\gamma)^{1/\gamma}, (2\kappa/\sin \theta\gamma)^{1/\gamma} \cdots$ from the origin. If we let $w = z^\gamma$ and $f(z) = g(w)$ then the real points of $g(w)$ on a line $\phi = \gamma\theta =$ constant are at distances $\kappa/\sin \theta\gamma, 2\kappa/\sin \theta\gamma, \cdots$ from the origin, showing that the real points lie on parallels to the x axis at distances $\kappa, 2\kappa, \cdots$ from it.

For purely imaginary points of $f(z)$ the phase must be an odd multiple of $\pi/2$. $\therefore \pi x^\gamma \phi(x) \operatorname{cosec} \pi\rho \sin \theta\gamma \sim (2\mu + 1)\pi/2$. If

$$\phi(x) \equiv 1, \; x^\gamma \sin \theta\gamma \sim (2\mu + 1)\kappa/2$$

and making the same transformation as before, we find that the imaginary points lie on parallels to the x axis at distances $\kappa/2, 3\kappa/2 \cdots$ from it.

Heretofore we have considered only entire functions whose zeros lie along one line. In this case from a knowledge of $n(r)$ we were able to obtain exact knowledge of the growth of $f(z)$ along any line. If however we assume the roots to lie along two distinct lines our results about the growth of $f(z)$ are not as precise as before. We shall see that this is an inherent difficulty of the problem.

Let

$$f(z) = \prod_{n=1}^{\infty} E\left(\frac{z}{a_n}, m\right) E\left(\frac{z}{b_n e^{i\theta}}, m\right)$$

where the a_n, b_n are real positive numbers. $f(z)$ has its zeros on the negative real axis and on the line making an angle θ with the positive real axis. If we know the number of $a_n \leqq r$ and $b_n \leqq r$ separately then we can consider

$$\phi_1(z) = \prod_{n=1}^{\infty} E\left(\frac{z}{a_n}, m\right), \qquad \phi_2(z) = \prod_{n=1}^{\infty} E\left(\frac{z}{b_n e^{i\theta}}, m\right)$$

independently and since $f = \phi_1 \cdot \phi_2$ we can obtain exact knowledge of the growth of $f(z)$. But in most cases all we know is the number of zeros inside a circle of radius r. This knowledge is very vague. It may happen that there are only a finite number of zeros on the negative real axis (which of course will not materially influence the growth of $f(z)$ or, on the other hand, almost all the zeros may be on the negative real axis, in which case it behaves just as do the functions of Theorem I.

Let us consider

$$f_1(z) = \prod_{n=1}^{\infty} E\left(\frac{z}{a_n}, m\right) E\left(\frac{z}{b_n}, m\right),$$

$$f_2(z) = \prod_{n=1}^{\infty} E\left(\frac{z}{a_n e^{i\theta}}, m\right) E\left(\frac{z}{b_n e^{i\theta}}, m\right).$$

If $z = re^{i\varphi}$ we can easily show that after neglecting a finite number of terms we will have

$$\log |f_2(z)| \geqq \log |f(z)| \geqq \log |f_1(z)|,$$

or

$$\log |f_2(z)| \leqq \log |f(z)| \leqq \log |f_1(z)|,$$

according as

$$\cos \gamma(\varphi - \theta) > \text{or} < \cos \gamma\phi.$$

Similar arguments can be used when the roots lie on more than two lines. We thus obtain the following theorem:

THEOREM II. *Let $f(z)$ be an entire function such that the number of its zeros inside a circle of radius $r = n(r) \sim Ax^\gamma\phi(x)$. Let the arguments of the zeros be $\theta_1, \theta_2, \cdots \theta_n$. (This is excessive. All that is needed is that the arguments approach θ_i as a limit). Then*

$$\log |f(re^{i\theta})| \sim (-1)^m \pi A \operatorname{cosec} \pi\rho K_1 x^\gamma \phi(x),$$

where K_1 is a real number included between the maximum and minimum of

$$\cos \gamma(\phi - \theta_1), \cos \gamma(\phi - \theta_2), \cdots \cos \gamma(\phi - \theta_n).$$

As mentioned previously, this is the best possible result unless we know more about the distribution of the zeros with the argument θ_i.

This theorem shows that even if the function behaves at ∞ exactly as the functions discussed in Theorem I, all we can say is that the arguments of its zeros approach π as a limit. Necessary and sufficient conditions that all the zeros of an entire function in a certain class be real have been given by Professor Wiener.[4] We will give a slight generalization of his results.

THEOREM III. *Let* $\psi(z) = \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{a_n^2}\right)$ *and let* $n(r)$ *be the number of* $| a_n |$ *not exceeding* r.

If $n(r) \sim Ar^\rho \phi(r)$, $0 < \rho < 1$, *then the necessary and sufficient condition for all the* a_n *to be real is that*

$$\lim_{y \to \infty} y^{1-\rho} [\phi(y)]^{-1} \int_{-y}^{y} x^{-2} \log | \psi(x) | \, dx = - \pi A \tan (\rho - 1) \frac{\pi}{2} \Big/ (\rho - 1)$$

On the assumption that the a_n are real

$$\frac{y^{1-\rho}}{\phi(y)} \int_{-y}^{y} x^{-2} \log | \psi(x) | \, dx = \frac{2}{y^\rho \phi(y)} \int_0^\infty dn(t) \frac{y}{t} \int_0^{y/t} s^{-2} \log | 1 - s^2 | \, ds$$

Let

$$K(\lambda) = \frac{1}{\lambda} \int_0^{1/\lambda} s^{-2} \log | 1 - s^2 | \, ds.$$

$$\int_0^\infty K(\lambda) \lambda^{i\alpha+\rho-1} \, d\lambda = - \pi \frac{\tan \frac{i\alpha + \rho - 1}{2} \pi}{i\alpha + \rho - 1} \neq 0, \quad (- \infty < \alpha < \infty)$$

Since

$$\lim_{y \to \infty} \frac{1}{y^\rho \phi(y)} \int_0^\infty G\left(\frac{t}{y}\right) dn(t) = \rho A \int_0^\infty G(t) \, t^{\rho-1} \, dt, \text{ we have}$$

$$\lim_{y \to \infty} \frac{y^{1-\rho}}{\phi(y)} \int_{-y}^{y} x^{-2} \log | \psi(x) | \, dx = - \pi A \tan (\rho - 1) \frac{\pi}{2} \Big/ (\rho - 1).$$

This proves the necessity.

[4] Paley and Wiener, ibid., Theorem XXIII.

Since

$$\int_0^\infty x^{-2}\left[\log\left|1-\frac{x^2}{a_n^2}\right|-\log\left|1-\frac{x^2}{|a_n|^2}\right|\right]dx>0,$$

the condition is also sufficient.

The theorem is also valid if $\phi(x)\equiv 1$ and $\rho=1$. Incidentally this theorem enables us to evaluate some indefinite integrals, e.g.

$$\int_0^\infty \frac{\log|\cos x|}{x^2}\,dx=-\frac{\pi}{2}=\int_0^\infty \frac{\log|J_0(x)|}{x^2}\,dx.$$

Exactly similar results hold for meromorphic functions. Instead of Theorem I we have

THEOREM IV. *Let* $g(z)=\sum\left[\frac{1}{a_n+z}-\frac{1}{a_n}+\cdots(-1)^m\frac{z^{m-1}}{a_n^m}\right]$ *where the* a_n *are real positive numbers such that* $0<a_1\leqq a_2\leqq\cdots$ *and* $\sum_{n=1}^{\infty}a_n^{-m-1}<\infty$. *Let* $n(r)$ *be the number of* a_n *not exceeding* r. *Let* $\lambda,\rho,\theta,\gamma$ *be real numbers such that* $\lambda>0$, $0<\rho<1$, $|\theta|<\pi$, $\gamma=m+\rho$. *Then the following statements*

i. $n(x)\sim\lambda x^\gamma\phi(x)$,

ii. $g(x)\sim(-1)^m\pi\lambda\gamma\operatorname{cosec}\pi\rho\,x^{\gamma-1}\phi(x)$,

iii. $g(xe^{i\theta})\sim(-1)^m\pi\lambda\gamma\operatorname{cosec}\pi\rho\,e^{i\theta(\gamma-1)}x^{\gamma-1}\phi(x)$,

are equivalent.

i. can be written

$$\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\int_0^\infty G\left(\frac{t}{x}\right)dn(t)=\lambda\gamma\int_0^\infty G(t)\,t^{\gamma-1}\,dt.$$

ii. can be written

$$\lim_{x\to\infty}\frac{g(x)}{x^{\gamma-1}\phi(x)}=\lim_{x\to\infty}\frac{1}{x^\gamma\phi(x)}\int_0^\infty\frac{(-1)^m\,dn(t)}{\left(\frac{t}{x}\right)^m\left(\frac{t}{x}+1\right)}$$

By separating iii. into real and imaginary parts we have

$$\lim_{x\to\infty}\frac{R[g(x)]}{x^{\gamma-1}\phi(x)}=\lim_{x\to\infty}\frac{(-1)^m}{x^\gamma\phi(x)}\int_0^\infty\frac{\cos(m-1)\theta+\frac{t}{x}\cos m\theta}{\left(\frac{t}{x}\right)^m\left[1+\frac{2t}{x}\cos\theta+\frac{t^2}{x^2}\right]}\,dn(t),$$

$$\lim_{x\to\infty}\frac{I\,[g(x)]}{x^{\gamma-1}\phi(x)}=\lim_{x\to\infty}\frac{(-1)^m}{x^{\gamma}\phi(x)}\int_0^{\infty}\frac{\sin\,(m-1)\,\theta+\dfrac{t}{x}\sin m\theta}{\left(\dfrac{t}{x}\right)^m\left[1+\dfrac{2t}{x}\cos\theta+\dfrac{t^2}{x^2}\right]}\,dn(t)$$

An application of Lemmas I, II and III completes the proof.

If we let $f(z)=\prod_{n=1}^{\infty}E\left(\frac{z}{a_n},\,m\right)$ then $g(z)=f'(z)/f(z)$ and our theorem shows that $f'(z)$ has the same order as $f(z)$. If we put $\phi(z)\equiv 1$ and $w=z^{\gamma}$, $F(w)=f(z)$, then $F'(w)=f'(z)/\gamma z^{\gamma-1}$. $\therefore F'(w)/F(w)\sim(-1)^m\,\pi\lambda$ cosec $\lambda\rho$ showing that $|\,F'(w)/F(w)\,|$ approaches the same finite limit along any line through the origin, excepting of course the lines on which $F(w)$ vanishes. For example.

$$\lim_{z\to\infty}\left|\frac{J_{\nu+1}(ze^{i\theta})}{J_{\nu}(z)}\right|=1,\qquad\qquad\theta\neq 0,\,\theta\neq\pi.$$

REACTION VELOCITY IN THE SYSTEM: $SOLID_I + GAS \rightarrow SOLID_{II}$*

By George S. Gardner

In the "dry purification" of illuminating gas, hydrogen sulphide is removed by passing the gas through beds of ferric oxide, forming a complex sulphide of iron. Oxygen acts upon this complex, regenerating ferric oxide. The following paper is intended to be preliminary to a study of reaction velocities in these systems.

Langmuir[1] has shown that the velocity of the typical catalytic solid-gas reactions is not limited by the rate of diffusion through an adsorbed film, but by the rate at which the molecules strike that portion of the surface which is active. However, he recognizes the fact that there are many chemical processes which are actually limited by physical factors, such as the rate of diffusion through layers of gas or moderately thick films covering the solid substances.

The object of the present investigation is to study the reaction velocity between a granular solid and a gas, for the case where the product of the reaction is a second solid, and the actual rate of reaction of the gas and solid at the reacting surface is great compared to the rate at which the reactant gas reaches the reacting surface. Many practical applications of such reactions deal with flowing gases; the present study, however, will be limited to reactions in closed systems.

Perhaps the earliest expression of reaction velocity between two phases from a diffusional standpoint was that of Nernst[2] and Brunner,[3] who studied the velocity of solution of magnesium oxide in acids. Based on the findings of Noyes and Whitney[4] as to the extremely rapid establishment of equilibrium at the boundary of two phases, Nernst expressed the amount of acid diffusing to the reacting surface in time dt as $DF\frac{c}{\partial}dt$, where D is the coefficient of diffusion, F the surface of the magnesia, c the concentration of acid, and ∂ the constant thickness of the film adhering to the magnesia surface (using his own nomenclature).

* Contribution from the Laboratory of Brewer and Gardner, Philadelphia, Pa.

[1] Langmuir, J. Am. Chem. Soc., 38, 2287 (1916).

[2] Nernst, Z. Physik. Chem., 47, 52 (1904).

[3] Brunner, Z. Physik. Chem., 47, 56 (1904).

[4] Noyes and Whitney, Z. physik. Chem., 23, 689 (1897).

Shortly following this classic work of Nernst and Brunner, Hantzsch,[5] neglecting the effect of diffusion, attempted to study the velocity of solid-gas reactions by applying the equations of homogeneous reactions. He made a considerable number of experiments upon the absorption of anhydrous ammonia and of hydrogen chloride by organic acids and bases. The usual procedure was to rub up one millimol of the solid absorbent with a small quantity of glass powder in an agate mortar. The absorption was made under varying conditions of pressure in three different types of apparatus. While the results did not show good agreement with his theoretical treatment, his data is of considerable value.

Predwoditelew and Witt[6] have made some careful measurements of the rate of reaction of pure anhydrous ammonia at constant pressure with anhydrous copper sulphate, on a micro scale. They used approximately 2 mg. of the solid reactant, and followed the course of the reaction by means of a micro-balance of special construction, working at pressures from 66 to 217 mm. Hg. Their theoretical treatment is believed by the writer to be in error on several points, one of which is that they attempt to avoid the effect of diffusion by using fine powders.

Considerable work of value on reactions of this nature has been done by several investigators on the corrosion of metals by various gases. The fact that the rate of increase of thickness of the reaction products is inversely proportional to the thickness of the reaction products at a given time seems to be well substantiated, and leads to the equation $w^2 = kt$, (where w is the amount of corrosion product formed in time t), if we take the thickness of the corrosion layer as a small fraction of the thickness of the corroding particle. The researches of Dunn,[7] Feichnecht,[8] Pilling and Bedworth,[9] and Tammann,[10, 11, 12, 13, 14] should be cited in this connection.

Aono,[15] by assuming that the rate of increase of thickness of reaction

[5] Hantzsch, Z. physik. Chem., 48, 289 (1904).
[6] Predwoditelew and Witt, Z. physik. Chem., 132, 47–54 (1928).
[7] Dunn, Proc. Royal Soc., 111, 210 (1926).
[8] Feichnecht, Z. Elekt., 35, 142 (1929).
[9] Pilling and Bedworth, J. Inst. Metals, 29, 529 (1923).
[10] Tammann, Z. anorg. allgem. Chem., 111, 78 (1920).
[11] Tammann, Z. anorg. allgem. Chem., 123, 196–224 (1922).
[12] Tammann, Z. anorg. allgem. Chem., 124, 25–35 (1922).
[13] Tammann, Z. anorg. allgem. Chem., 128, 179–206 (1923).
[14] Tammann, Rec. trav. chim., 42, 547–51 (1923).
[15] Aono, Bull. Chem. Soc. Japan, 7, 274–80 (1932).

product layer is either inversely proportional to the thickness, or is constant, has expressed the velocity of azotation of calcium carbide by satisfactory equations. He has calculated the thickness of the layer of reaction products as a function of the amount of solid reactant formed, in a similar manner to the method used by the writer.

DEVELOPMENT OF THE EQUATION FOR REACTION VELOCITY

Let us then visualize a closed system consisting of a granular solid and a mixture of gases, at substantially constant temperature. Let one of the gases react chemically with the solid reactant to form a second solid, and assume that the reactant gas is present in sufficient quantity to react with all of the solid present in the active form. Such a system might be the absorption of gaseous ammonia by solid copper sulphate to form a complex compound, as mentioned above. We will assume that the particles of the solid are approximately spherical, homogeneous, of nearly the same size, and are each accessible to the reactant gas to approximately the same degree. Let the number of particles be n.

If we confine our attention to one of the spheres of solid reactant, and try to understand the mechanism of such a process, we see that there are three questions to be answered if we are to be able to express the velocity of a reaction of this nature mathematically:

1. What is the extent of the effective reacting surface?
2. What is the resistance to diffusion?
3. What is the intensity of the driving force that pushes the diffusing gas across this resistance?

In order to answer the first question, we must refer to the work of Langmuir.[16] He has shown that when a solid reacts with a gas to form a second solid, if the phase rule indicates the simultaneous existence of two solid phases, it is necessary that the reaction shall occur only at the boundary between the two phases,—that is, that the rate of the reaction should be proportional to the extent of the boundary between the two solid phases. According to this theory, there should be no reaction between a gas and a solid, if there were no trace of the second solid phase present. However, he assumes that in this case a gas would be slowly adsorbed to form a solid solution. Gradually nuclei would develop in this solid solution and the second solid phase would appear. The rate of reaction would then increase as the surface area of the second phase grows. This initial slowness to react may be referred to as the "induction period."

[16] Langmuir, J. Am. Chem. Soc., 38, 2263 (1916).

In the reactions which we shall study here, this initial adsorption is rapid, so that the surface of all the solid granules will be very quickly covered with a layer of the second phase. We assume, however, that the initial adsorption reaction on the pure solid reactant phase must be slow compared to the reaction at the $solid_I$–$solid_{II}$ phase boundary.

When dealing with a granular solid having an appreciable induction period, it does not seem likely that it would be possible to express its reaction velocity over the whole range by any simple means, as it is evident that during the initial period some of the granules would be reacting and others would not. An example of this is the reduction of cupric oxide, which has been studied by Pease and Taylor.[17] "Reduction commenced visibly by the formation of a red speck amongst the little mass of copper oxide at the lowest part of the bulb; this quickly spread, becoming rapidly larger and larger until practically the whole mass became a reddish metallic powder, the progress of the reaction reminding one of the slow deflagration of a quantity of touch paper, the reducing action being apparently propagated from particle to particle through the mass. This occurred, be it remembered, in the presence of an excess of hydrogen." Thus, as soon as the induction period has ended, which is to say that all the solid granules are covered with the initial thin layer of solid reaction product,—which is almost instantly in many reactions of this type,—we may begin to form some estimate of the extent of the phase boundary, or what is to say the same thing, the extent of the reacting surface. If, then, we consider a single granule of solid absorbent, the area of the phase boundary is very evidently proportional to the area separating the inner portion of unreacted solid from the outer shell of reaction product.

The answer to the second question is that the resistance to diffusion is given by the thickness of this shell of reaction products mentioned above, plus the constant resistance offered by any gaseous, liquid, or solid film which might be present on the surface of the solid granules. For gases in closed systems, any film present would be of the nature of an adsorbed film. All of the work of Langmuir has pointed toward the fact that this adsorbed film of gas is not likely to exceed one molecule in thickness, hence we may perhaps disregard it here. However, this does not exclude solid or liquid films. The resistance to diffusion is thus the resistance of the solid reaction product, plus the resistance of any film, and we make no assumption as to the nature of this film, other than constant diffusional resistance. We shall write v as the variable

[17] Pease and Taylor, J. Am. Chem. Soc., 43, 2179 (1921).

resistance of the solid reaction product at time θ, and g as the constant resistance of the film, expressed in similar units to that of v.

In answer to our third question, we may say that by analogy to the flow of heat, the driving force causing diffusion is the overall pressure gradient of the reactant gas from the external surface of the absorbent granule to the reacting surface. Then we may write

$$\phi = (p - p_s)/(v + g),$$

where

ϕ = pressure gradient,
v = resistance of layer of reaction product,
g = resistance of film,
p = partial pressure of the reactant gas at the surface of granule,
p_s = partial pressure of the reactant gas at the phase boundary.

Since we have assumed the actual reaction rate to be great compared to the rate of supply of reactant gas to the phase boundary, p_s may be taken to be negligible compared to p, for reactions in which the pressures are not extremely small, and the relation becomes

$$\phi = p/(v + g).$$

By further analogy to the conduction of heat, the instantaneous rate of transfer of reactant gas to the phase boundary (reacting surface) is given by:

$$dw/d\theta = k_1\phi s,$$

where w is the mass of gas striking active surface, s is the area of the boundary between the solid phases, and θ is the time.

Since the mass dw of reactant gas reaching the phase boundary is equal to the amount dz that reacts, we have, inserting our value for ϕ,

$$dz/d\theta = k_1ps/(v + g), \qquad (1)$$

which says that the amount of solid reaction product formed on one sphere of solid absorbent in unit time is proportional to the gas pressure, the area of the boundary between the solid phases, and inversely proportional to the combined resistance of the layer of reaction product and the film.

It is evidently impractical to make direct measurements of such quantities as the thickness of a layer of reaction product at a definite time, the thickness of a solid, liquid or gaseous film, and of the area of

the phase boundary at that time. However, by means of an indirect approach to the problem, we may make considerable progress toward such measurements. If the sphere of absorbent is exposed to the reactant gas for an indefinitely long time, it will absorb a certain quantity of the gas z_∞. We may say that the absorption capacity of the solid at $\theta = 0$ is equal to this quantity. We shall write this absorption capacity at time $\theta = 0$ as y_0 and at any other time as y. (y will thus equal $(z_\infty - z.)$ If the solid is homogeneous, this absorption capacity y varies directly as the volume of the unreacted solid. Let us then set this volume equal to $k_2 y$. With this view of the problem, we may readily define the area of the phase boundary and the thickness of the layer of reaction products. The former becomes $k_3 y^{2/3}$ and the latter $k_4(y_0^{1/3} - y^{1/3})$. The amount of solid reactant formed is $k_5(y_0 - y)$. This development is based on the reaction product occupying approximately the same volume as the reactant.[18]

Rewriting equation (1) in the light of the above,

$$\frac{d(y_0 - y)}{d\theta} = \frac{(k_1 k_3/k_4 k_5) y^{2/3} p}{(y_0^{1/3} - y^{1/3}) + g/k_4} \tag{2}$$

It is evident that the variation of $(y_0^{1/3} - y^{1/3})$ with $(y_0 - y)$ is given by

$$\frac{d(y_0^{1/3} - y^{1/3})}{d(y_0 - y)} = \frac{1}{3y^{2/3}}.$$

Combining this with equation (2) we get

$$\frac{d(y_0^{1/3} - y^{1/3})}{d\theta} = \frac{(1/3)(k_1 k_3/k_4 k_5) p}{(y_0^{1/3} - y^{1/3}) + g/k_4},$$

which holds for one sphere of solid absorbent. For n spheres of absorbent the above equation will contain the additional constant n in

[18] It is interesting to consider at this point the consequences of the reaction product occupying a decidedly different volume from the unreacted solid. Since we have put y proportional to the volume of the unreacted solid, there will be no error in the value of the surface $k_3 y^{2/3}$, and since the ratio of the volume of the reaction product to the volume of the reactant is a constant, $k_4(y_0^{1/3} - y^{1/3})$ is affected only by a constant factor, which will not affect our equations. $k_5(y_0 - y)$ will in the same way be affected only by a constant. If the swelling or shrinking effect were great, we should expect to rupture the sphere, which would of itself prevent the application of a mathematical treatment. In addition, it seems possible that a considerable variation in the size of the absorbing sphere might result in a change in the film thickness. Such an effect is beyond the scope of this paper.

the right hand member. With this in mind, we may rewrite the equation as

$$\frac{d(y_0^{1/3} - y^{1/3})}{d\theta} = \frac{Ap}{(y_0^{1/3} - y^{1/3}) + B}, \tag{3}$$

where A and B represent groups of constants. Equation (3) represents the generalized equation for reaction velocity in the systems which we are considering.

INTEGRATION OF EQUATION (3) FOR SPECIAL CASES

We may find that under the conditions of an experiment equation (3) may assume simplified forms. The resistance of the solid reaction product to the diffusion of the gas toward the reacting surface may be very large in comparison to the resistance of any film present on the surface of the solid granule. Again, the resistance of the reaction product may be negligible compared to the resistance of such a film. In a third case we might find that both of these resistances are essential in expressing the time course of the reaction. The following is a short summary of some of these special cases, with their integration, where practical:

Case A. Reaction at constant pressure of reactant gas, Film controlling.

In this case, equation (3) takes the simplified form:

$$d(y_0^{1/3} - y^{1/3})/d\theta = (A/B)p,$$

since $(y_0^{1/3} - y^{1/3})$ is very small compared to B. Integrated,

$$(y_0^{1/3} - y^{1/3}) = (A/B)p\theta + \text{constant}, \tag{4}$$

where $(A/B)p$ is constant.

Case B. Reaction at constant pressure of reactant gas, Reaction product controlling.

For this case, we may write equation (3) as

$$d(y_0^{1/3} - y^{1/3})/d\theta = Ap/(y_0^{1/3} - y^{1/3}),$$

since B is very small compared to $(y_0^{1/3} - y^{1/3})$. Integration would lead to

$$(y_0^{1/3} - y^{1/3})^2 = (2Ap)\theta + \text{constant},$$

where the term $(2Ap)$ is constant.

Case C. Reaction at constant pressure of reactant gas,
Both reaction product and film controlling.

If we integrate equation (3) as written, taking p as constant, with the boundary condition that $y = y_0$ when $\theta = 0$, and divide the integrated form by $(y_0^{1/3} - y^{1/3})$, we get

$$(y_0^{1/3} - y^{1/3}) = [(2Ap)\theta/(y_0^{1/3} - y^{1/3})] - 2B \tag{6}$$

Case D. Variable pressure of reactant gas,
Film controlling.

Equation (3) becomes

$$d(y_0^{1/3} - y^{1/3})/d\theta = (A/B)p$$

under these conditions, where p is now a variable.

Frequently we may express y as a function of p by means of the following procedure:

$$y_0 = a(p_0 - p_\infty),$$

i.e., the absorption capacity at time $\theta = 0$ is proportional to the total drop in pressure in the system between $\theta = 0$ and $\theta = \infty$, and

$$y = a(p - p_\infty),$$

i.e., the absorption capacity at time θ is proportional to the difference of pressure between time θ and time θ_∞. Thus we may substitute these values in the above equation, obtaining

$$\frac{da^{1/3}[(p_0 - p_\infty)^{1/3} - (p - p_\infty)^{1/3}]}{d\theta} = (A/B)p$$

For the special case in which equivalent quantities of reactant gas and solid have been taken ($p_0 = 1$ and $p_\infty = 0$),

$-a^{1/3}\,dp^{1/3} = (A/B)p\,d\theta$, which, when integrated, gives:

$(a^{1/3}B/2A)(1/p^{2/3}) = \theta +$ constant.

Case E. Variable pressure of reactant gas,
Reacted layer controlling.

For this case,

$$d(y_0^{1/3} - y^{1/3})/d\theta = Ap/(y_0^{1/3} - y^{1/3}).$$

Upon substitution of the above values for y in terms of p, and integration between appropriate limits, we have:

$$\frac{a^{2/3}}{2}\Big|_{p_0}^{p}[(p_0 - p_\infty)^{1/3} - (p - p_\infty)^{1/3}]^2 = \int_{\theta_0}^{\theta} Ap\, d\theta$$

Perhaps the simplest method to utilize in the solution of this equation is to evaluate the integral on the right hand side graphically, and then plot this result against the value of the left hand side of the equation.

Case F. Variable pressure of reactant gas,
Both reacted layer and film controlling.

A graphical solution, similar to the above, is suggested for this case.

APPLICATION OF REACTION VELOCITY EQUATIONS TO DATA RECORDED IN LITERATURE

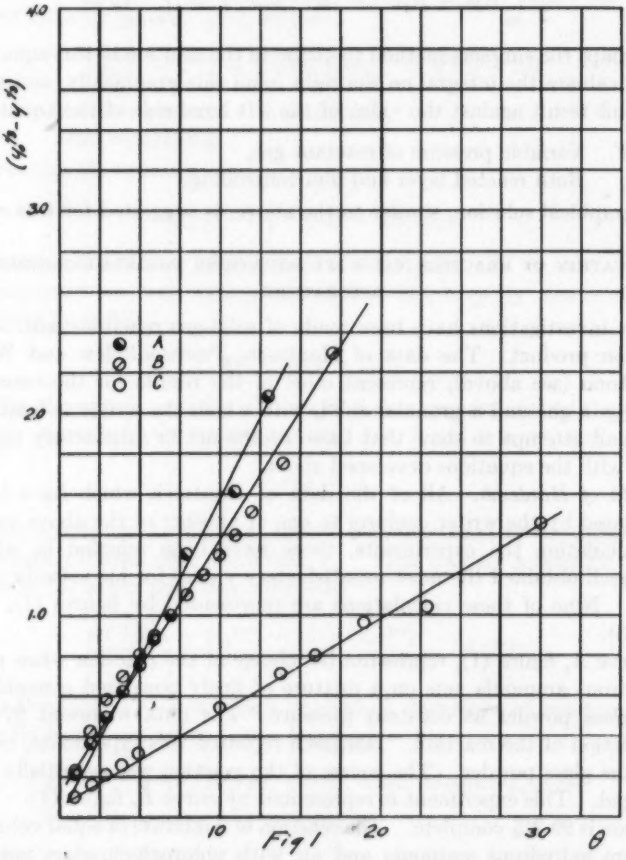
Few investigations have been made of solid-gas reactions with solid reaction product. The data of Hantzsch, Predwoditelew and Witt, and Aono (see above), represent most of the results for the reaction between a gas and a granular solid, with which the writer is familiar. We shall attempt to show that these results are in satisfactory agreement with the equations developed above.

Data of Hantzsch. All of the data of Hantzsch which have been calculated by the writer conform to one or another of the above cases. In calculating the experiments, those have been selected in which Hantzsch obtained the most unsatisfactory values for his velocity constant. Nine of these calculations are represented by figures (1), (2), and (3).

Curve *A*, figure (1), represents the course of the reaction when pure anhydrous ammonia acts on a mixture of finely powdered α naphthol and glass powder at constant pressure. The data represent 97.9% completion of the reaction. Hantzsch repeated this experiment, omitting the glass powder. The course of the reaction was essentially unchanged. This experiment is represented by curve *B*, figure (1). The reaction is 99.2% complete. The reaction of a mixture of equal volumes of pure anhydrous ammonia and air with phloroglucin-glass powder mixture at constant pressure of reactant gas is shown in curve *C*, figure (1). The reaction was taken to 78.7% of completion. In all of these experiments the velocity is expressed by equation (4), which is consistent with a film-controlled mechanism.

Curve A, figure (2), shows the course of the reaction when pure an-

hydrous ammonia and phthallic acid-glass powder mixture react at constant pressure. Reaction is 96.2% complete. Curve *B*, figure (2), represents the velocity of the reaction of pure anhydrous ammonia on



a finely powdered mixture of tartaric acid and glass powder at constant pressure. The reaction is 79.0% complete. Curve *C*, figure (2) represents the reaction of pure anhydrous ammonia with a finely powdered mixture of hydroquinone and glass powder at constant pressure. The reaction is 75.3% complete. Curve *D*, figure (2), represents the reac-

tion of a 50% anhydrous ammonia-air mixture, with finely powdered β naphthol and glass powder, at constant pressure of reactant gas. The reaction is 67.7% complete. The velocity of these four reactions is

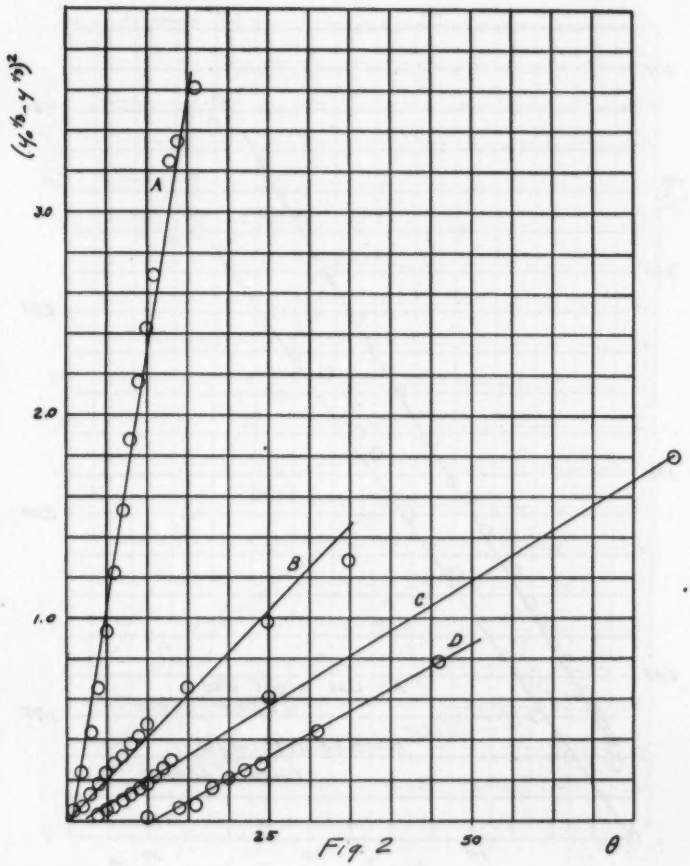


Fig 2

expressed by equation (5), which suggests that the reacted-solid layer controls the reaction rate.

Curves *A* and *B*, figure (3), represent the reaction of *p*-anisidine in 48% hydrochloric acid—52% air mixture and pseudo-cumenol in 66%

ammonia—34% air mixture, respectively. The reactions are 63.5% and 51.9% complete. In each case equivalent quantities of the solid and reactant gas were taken; hence the pressure of the reactant gas

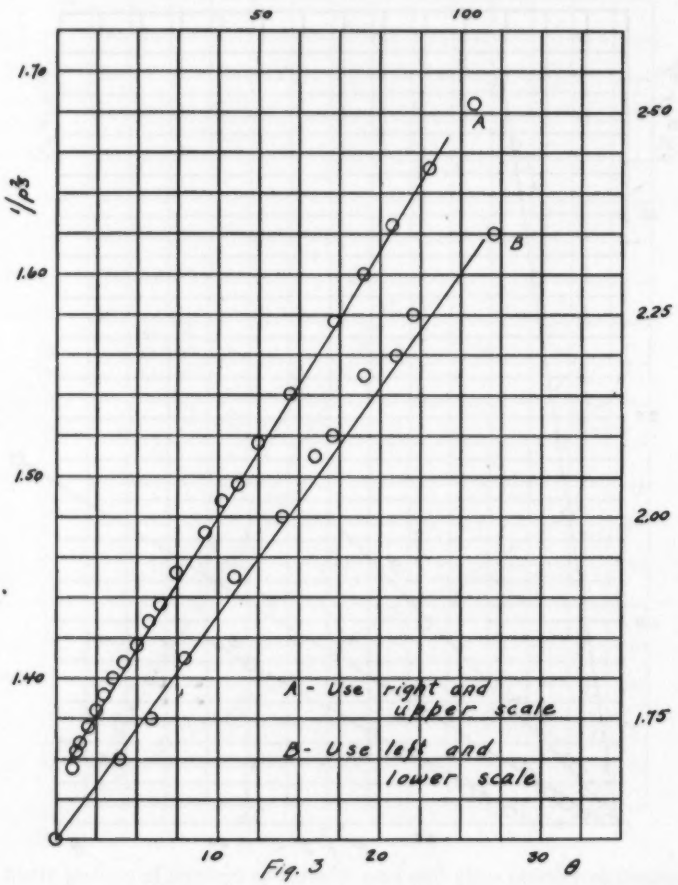


Fig. 3

varies. The formulation corresponds to equation (7), which is in agreement with the hypothesis that the film is here the controlling resistance.

In the curves expressing Hantzsch's data at constant pressure, y represents cc. of reactant gas; for the data at variable pressure, p represents

pressure of the reactant gas in atmospheres. Since he did not always carry his reactions to completion, we have given the per cent of completeness for each experiment calculated. The method used by him for the introduction of the solid into the reactant gas,—introduction under

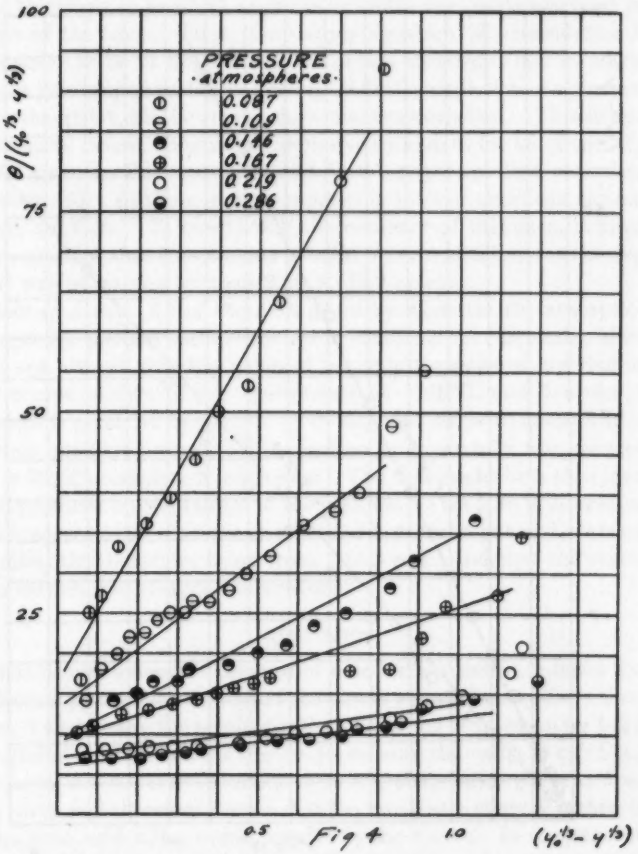


Fig 4

the surface of mercury into the reaction tube,—is open to the objection that the solid contained some entrained air when it was introduced into the gas space. Probably this was responsible for some irregularities noted in his results at small values of the time, (figure 1, curve *B* and

figure (3), curve *A*). Also, it is worth noting that in the experiment in which the glass powder was omitted, the curve was slightly irregular. The effect of the glass powder seems to be to make the solid granules more readily accessible to the gas.

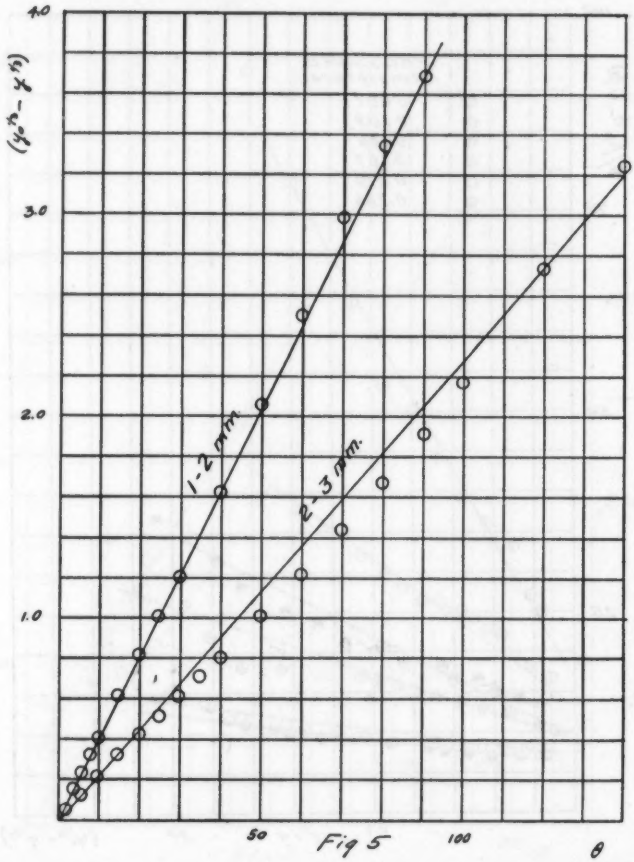


Fig 5

Data of Predwoditelew and Witt. We might feel that the data of these workers on the absorption of ammonia by anhydrous copper sulphate should conform to equation (4) or (5), since the experiments were made at constant pressure, and we should expect either the film or the re-

acted layer to be controlling. However, although equation (5) gives a straight line over a limited time interval, it does not express the rapid absorption at the beginning of the reaction with any degree of satisfaction. This leads to the thought that there must be a film of sufficient thickness to exert an appreciable resistance, particularly at small values of the time. Thus, if we apply equation (6), considering both resistances to be of importance, we obtain a straight line in all cases, up to a point representing an average of 88% completion of the reaction. After this point, all the curves show marked deviation. This deviation is discussed below. Figure (4) represents the data for all of the experiments given in Predwoditelew and Witt's paper, plotted according to equation (6). In these plots, y represents mols of ammonia per mol of copper sulphate. In considering the accuracy of this data, it must be borne in mind that the amount of solid absorbent taken for the experiment was never greater than 2.55×10^{-3} grams.

Data of Aono. All of the data given by Aono for the absorption of nitrogen by calcium carbide are not in condition for calculation since he does not always give the value of y_0, or any means of arriving at it. The course of two of his experiments at 1140°C. and 1 atmosphere pressure is given by figure (5). y in this plot refers to amount of solid reaction product formed. The reaction is apparently film controlled, and is 99.6% complete in each case. The film might be a thin layer of fused products on the surface of the granule. This data is interesting in that it serves as an example of reactions between a gas and coarse solid granules, the diameters being from 1 to 3 mm., and that the reactions were carried out at high temperatures.

DISCUSSION

That completeness in this type of reaction is reached in finite time is evidenced by the fact that all of the above equations are discontinuous. Thus, if we observe the results for the absorption of nitrogen by 1–2 mm. granules of calcium carbide (figure 5), we note the value of $(y_0^{1/3} - y^{1/3})$ at $\theta = 90$ is 3.69, corresponding to 99.6% of the absorption at $\theta = \infty$. The curve will therefore show a definite break at a slightly higher value of the time, and will proceed parallel to the θ axis. In the experiment cited, the break comes at $(y_0^{1/3} - y^{1/3}) = 4.10$ (not shown in the figure).

If we observe equation (3) for a film-controlling reaction at constant pressure, the form of the equation would lead us to the conclusion that the increase of $(y_0^{1/3} - y^{1/3})$ with θ would go on indefinitely at a constant rate. However, when we consider what takes place when the last bit

of the interior of the granule is reacted with, the matter assumes a different aspect; for, in fact, at this point the pressure of the reactant gas at the surface is not close to zero, but equal to p; and $(p - p_s)$, which is what we really mean when we use p in equation (3), becomes, not p, but zero.

A similar condition exists at the start of the reaction. For a reaction-product controlling reaction at constant pressure, at $\theta = 0$, the equation would seem to indicate infinite velocity. If we write $(p - p_s)$ for p, and take $p_s = p$ (which it is at time $\theta = 0$), the right hand side of the equation takes the indeterminate form 0/0.

This helps us to understand the curves of figure (4), results of Predwoditelew and Witt. The break comes at values varying from 80% of complete absorption at 0.087 atm. pressure, to 95.8% of complete absorption at 0.167 atm. pressure, and shows no definite relation to pressure, but seems to depend rather on the experimental conditions,—accuracy of the balance, etc. When we consider that these men worked with samples of approximately 2 mg. of solid, we realize that the accuracy attained by them shows remarkably fine technique.

If we consider a reaction whose velocity conforms to case *A* or *D* (film controlling), the question arises, how do we know that such a reaction is controlled by a process of diffusion? Is it not just as likely that the equation expresses the rate of reaction *at* the surface as it is that it expresses the rate of supply of reactant gas *to* the surface? If the differential equation takes the form

$$\frac{d(y_0^{1/3} - y^{1/3})}{d\theta} = \text{constant, or constant} \times \text{pressure,}$$

there is nothing in the equation itself to indicate that the constant is really composed of a component of film resistance. However, there are several reasons which point toward an actual film resistance. The first is that there are reactions which conform to case "C,"—both reaction product and film controlled. Another reason is the following: When a solid reaction product is formed it seems highly probably that it would impede the reactant gas molecules to some extent,—at least more than any resistance the gas molecules would meet in the free gas space, and it seems reasonable that this resistance of the solid reaction product would increase as the reaction product grows in thickness. Now, if, as the resistance of the reaction product varies, the reaction velocity equation still shows the presence of a constant resistance factor, then this latter must of necessity be much greater than the resistance of the solid

reaction product, and would indicate the presence of a resistant film. Thus, in equation (1), if $k_1/(v + g)$ remained practically constant, notwithstanding the fact that v is changing, the presence of the constant g in the denominator is essential.

The derivation given is intended to hold for all solid-gas reactions (with the limitations mentioned), in which the product of the reaction is a second solid, for which the phase rule would demand the existence of two solid phases, and not a solid solution. The results give a confirmation of Langmuir's assumption that for such conditions the reaction is centered at the solid_I–solid_II boundary. If a solid solution were formed, rather than two solid phases, the reaction would not predominate at a definite distance toward the inside of the absorbing sphere. If the former were the case, the center of the solid sphere would at all times act as the inner diffusional boundary, and the solid reaction product (the solid solution) would be distributed throughout the sphere according to some other law,—perhaps according to Fourier's law for the flow of heat in a solid sphere.

Most of the data cited refers to reactions between a gas and a fine powder. However, the work of Aono, some of whose results are given in figure (5), has been carried out with granules of appreciable size. In applying these equations to large granules, we must observe that if the granule is of loose construction it may act in effect as a number of smaller granules. However, if the accessibility of each small particle of the larger granules is the same, and the size is approximately the same, the net result should not be seriously affected. Some workers, among which we list Hantzsch and Predwoditelew and Witt, have attempted to avoid diffusion effects by the use of very fine powders. This would evidently be impossible unless the particles of the powder were of atomic dimensions.

It is not always possible to predict whether the reaction will follow a film-controlled or a reaction product—controlled course, merely from observing the physical characteristics of the solid. Experiment is the only safe guide in this respect at the present state of our knowledge. The results indicate that the mechanism for absorption from a pure gas is the same as that from a gas mixture, for closed systems. In flowing gases, of course, we usually have a decided gas film to deal with.

Although the above equations have been derived for spheres, a similar derivation for other solid shapes might be a simple matter. For example, for absorption by flat slabs, i.e., plate-like crystals, the formulation should be easy to work out.

The writer is of the opinion that an equation of the form of equation (1) might express the velocity of all solid-gas, diffusion-controlled reactions, for any form of reaction product,—solid, liquid or gaseous. Here the equation would assume the more general form

$$\frac{dz}{d\theta} = \frac{k_1 ps}{(u + v + \cdots) + (h + i + \cdots)},$$

where the term in the denominator represents the resistance to diffusion of the various constant and variable resistances. In the case of a liquid reaction product, we would have extremely complicated conditions, where the liquid would tend to vaporize, run off the solid surface, etc. It is almost superfluous to state that an immense amount of work needs to be done on solid-gas reactions before we shall be in a position to formulate them exactly.

The writer wishes to thank Drs. Frank L. Hitchcock of Massachusetts Institute of Technology, E. R. Weaver of the Bureau of Standards, and J. Edward Brewer, of this laboratory, for reading the manuscript and for helpful suggestions.

SUMMARY

1. Based on the published work of Langmuir, and Fourier's law for the conduction of heat, the equation

$$dz/d\theta = k_1 ps/(v + g),$$

has been derived to express the rate of a solid-gas reaction, where the product is a second solid, and the phase rule would demand the presence of two solid phases. z = amount of gas reacting, θ = time, p = pressure, s = area of boundary between two solid phases, v = resistance of solid reaction product to diffusion of reactant gas, g = resistance of film to diffusion of reactant gas. The limiting conditions are that the particles of the solid be approximately spherical (this condition seems to be satisfactorily attained in most powders), homogeneous, of nearly the same size, are each accessible to the reactant gas to approximately the same degree, and that the induction period of the reaction be short.

2. It has been shown that the area of the solid-phase boundary and the resistance of the solid reaction product can be expressed in terms of y, the absorption capacity at any time θ. y may be expressed in terms of the amount of solid reaction product formed, the amount of gas used up, or the pressure of the reactant gas.

3. The integration of the equation is performed or indicated for six special cases.

4. All of the data of this nature with which the writer is familiar have been found to conform to one or another of the integrated forms.

5. The results indicate:

a) It is impossible to avoid diffusion effects by using fine powders;

b) Absorption from a pure gas and from a gas mixture follow the same laws, in a closed system;

c) Completeness in reactions of the kind studied is reached in finite time;

d) A confirmation of Langmuir's assumption that for such reactions the gas reacts at the boundary of the two solid phases.

6. The general equation

$$\frac{dz}{d\theta} = \frac{k_1 ps}{(u + v + \cdots) + (h + i + \cdots)},$$

is proposed for the expression of the reaction velocity of any solid-gas, diffusion-controlled reaction, where the term in the denominator represents the resistance to diffusion of the various constant and variable resistances.

NOTE ON THE ABSORPTION SPECTRA OF OXYGEN AT HIGH CONCENTRATION

By B. J. Eiseman, Jr.

The results of an investigation of the absorption spectrum of oxygen, carried out in 1927 and 1928, have not hitherto been published.[1] More recent data[2] have superseded these results in large part. In this note are presented those observations which are still "new" and that part of the data which is comparable in accuracy to the more recent data.

The oxygen was prepared by electrolysis of pure dilute sulphuric acid and passed over calcium chloride, solid sodium hydroxide, platinized asbestos at 200°C. and phosphorous pentoxide in the order named. It was distilled four times, the first and last portions being rejected, in all about half the total quantity. Thus, the oxygen used was more carefully purified than that used in other investigations. That the absorption bands are due to an absorber composed of oxygen only follows from the purity of the oxygen.

The apparatus has already been described.[3]

The bands at λλ 6280, 5765 and 4775 Å., observed at room temperature at pressures up to 440 atm., were found to degrade toward the short wave-lengths. This is of interest in connection with the interpretation of the bands. Shading toward the short wave-lengths has been observed by others only in liquid oxygen.

The band at λ 4470 Å. appeared only at low temperature at constant concentration. The oxygen was at about 400 atm. pressure at room temperature and was observed at a temperature just above that of the disappearance of the liquid phase. This indicates an intensification of the absorption with decreasing temperature in accord with the not very certain conclusions of Liveing and Dewar.[4]

[1] Given in part in: Eiseman, Thesis, M. I. T., 1927.

[2] J. C. McLennan, H. D. Smith and J. O. Wilhelm, Trans. Roy. Soc. Can., **24**, (III) 65 (1930).
J. W. Ellis and H. O. Kneser, Z. für Physik, **86**, 583 (1933).
R. Guillien, Compt. rend., **198**, 1223 (1934).
W. Finkelnburg and W. Steiner, Z. für Physik, **79**, 69 (1932).

[3] B. J. Eiseman, Jr., J. Am. Chem. Soc., **54**, 1778 (1932).

[4] G. D. Liveing and J. Dewar, Proc. Roy. Soc., **46**, 222 (1889).

The wave-lengths of the ultraviolet "triple" bands are given in the following table.

Wave-lengths in Å.U. (in air)

Broad band center	Broad band limits	Narrow band maxima	Broad band center	Broad band limits	Narrow band maxima
				2565	2560
2847			2552		2551
				2540	2544
	2805	2797.5		2533	2527
2786		2784.5	2520		2519
	2767	2772		2508	2511
	2749	2741		2503	2498
2732		2729	2491		2490
	2715	2718		2480	2484
	2692	2685		2477	2473
2678		2675.5	2466		2465
	2664	2668		2455	2458
	2647	2642			
2633		2631	2444		
	2619	2622			
	2604	2598			
2590		2589	2434		
	2576	2580			

The narrow band maxima of absorption are accurate to about ±2 Å. The pressures range from 440 atm. at about 20°C. for the bands of longer wave-length to about 55 atm. for those of shorter wave-length.

A paper concerning the interpretation of these and other relevant data will be published elsewhere.

A PARTICULAR BOUNDARY VALUE PROBLEM[1]

By Maurice Julius Fish

Introduction

This thesis presents the solution of a certain boundary value problem involving a differential equation which arises in mathematical physics. The integrability conditions[2] of the shearing stresses in the torsion of a ring sector give rise to an elliptic partial differential equation of the second order;[3] namely,

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} - \frac{3}{x}\frac{\partial\phi}{\partial x} = k \tag{1}$$

The same differential equation arises in the analytic study of the effects of side leakage of oil in bearings,[4] and again, in the flow of electricity through a thin homogenous plate of variable thickness.[5]

In 1899, J. H. Michell[6] solved the torsion problem, where he used a rectangle as a contour. A. Timpe[7] obtained a solution in the form of an integral, where he considered a shaft of variable cross section, but of infinite length, subjected to torsion. L. S. Jacobson[5] made use of the electrical analogy of the problem and obtained an experimental solution by an investigation of the distribution of electrical potential in the thin

[1] This paper was presented in partial fulfillment for the degree of Doctor of Philosophy in mathematics at the Massachusetts Institute of Technology. The author wishes to thank his teacher, Professor Philip Franklin, for valuable assistance.

[2] These integrability conditions are known as compatibility equations in the theory of elasticity.

[3] Theory of Elasticity. S. Timoshenko, 1st ed., 1934, pp. 276 and 355. McGraw Hill Book Co.

[4] Effects of Side Leakage in 120-Degree Centrally Supported Journal Bearings. S. J. Needs, Trans. Amer. Soc. Mech. Eng., vol. 56, 1934, p. 721.

[5] Torsional Stress Concentrations in Shafts of Circular Cross Section and Variable Diameter. L. S. Jacobson. Trans. Am. Soc. Mech. Eng., vol. 45, 1925, p. 619.

[6] The Uniform Torsion and Flexure of Incomplete Tores, with Application to Helical Springs. J. H. Michell, Proc. London Math. Soc., vol. 31, 1899, p. 130.

[7] Die Torsion von Umdrehungskörpern. A. Timpe. Math. Annalen, vol. 71, 1911, p. 480.

plate. In 1931, O. Göhner[8] considered the ring sector under torsion and obtained a solution by the method of successive approximations.

From the theory of elliptic differential equations, the solution is unique,[9] and therefore, in the present paper, only the actual determination of the solution will be considered. The function $\phi = \phi(x, y)$ of equation (1) has the nature of a stress function,[10] and is taken as a constant on the boundary, where an annular ring, $r_1 \leqq r \leqq r_2; 0 \leqq \theta < 2\pi$, $r_2 - r_1 > 0$ is chosen for the boundary. The value k in (1) is a constant which depends upon the external twisting moment. It is to be noted that the origin of coördinates for the boundary chosen is inside the ring while equation (1) in the case of the ring sector under torsion was developed with the origin of coördinates outside the circle. Therefore, this paper can not be considered a solution of the ring sector under torsion, but the analysis given throws considerable light on the differential equation involved.

The method of solution employed here is first to reduce equation (1) to the homogeneous form with a change in the boundary conditions of ϕ. Then, the method of particular solutions of Daniel Bernoulli is used. After separation of variables, the radial equation presents no difficulties, but the angular equation which is the form

$$(1 - w^2)\frac{d^2\Theta}{dw^2} + 2w\frac{d\Theta}{dw} + m\Theta = 0$$

has more interest. Particular solutions of this latter equation are obtained in terms of sets of functions which may be orthogonalized over an interval. Making use of the values of the stress functions at the boundary, the constants of integration are determined by evaluating the Fourier coefficients.

Solution

The particular problem which is to be considered now may be stated as; it is required to find the solution $\phi(x, y)$ of the differential equation

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} - \frac{3}{x}\frac{\partial\phi}{\partial x} = k \qquad (1)$$

[8] Schubspannungsverteilung im Querschnitt eines gedrillten Ringstabs mit Anwendung auf Schraubenfedern. O. Göhner. Ingenieur-Archiv, vol. 2, 1931, p. 1.

[9] Cours d'Analyse Mathematique. E. Goursat, 4th ed., 1927, vol. 3, Gauthier-Villars et Cie, p. 224.

[10] Stress Functions. H. B. Phillips. Journal Math. and Phys., vol. 13, 1934, p. 421.

such that $\phi(x, y)$ vanishes on the boundary of the ring sector, $r_1 \leqq r \leqq r_2$; $0 \leqq \theta < 2\pi$ where $r_2 - r_1 > 0$.

For the solution, first, reduce equation (1) to the homogeneous form with an accompanying change of boundary conditions.
Let

$$\phi(x, y) = ax^2 + \psi(x, y), \qquad a = \text{const.} \tag{2}$$

Then

$$\phi_x = 2ax + \psi_x$$

and

$$\phi_{xx} = 2a + \psi_{xx}$$

Hence, equation (1) becomes

$$2a + \psi_{xx} + \psi_{yy} - \frac{3}{x}(2ax + \psi_x) = k$$

or

$$\psi_{xx} + \psi_{yy} - \frac{3}{x}\psi_x - 4a = k$$

If $k = -4a$ this equation reduces to the form

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} - \frac{3}{x}\frac{\partial\psi}{\partial x} = 0 \tag{3}$$

But since ϕ has been taken equal to zero on the boundary, the value of ψ on the boundary from (2) is $-(ax^2) = \frac{k}{4}x^2$.

Now, transform equation (3) into polar coordinates, as follows:
Let

$$x = r\cos\theta, \quad y = r\sin\theta, \quad \theta = \tan^{-1} y/x, \quad r^2 = x^2 + y^2$$

Then

$$\frac{\partial\psi}{\partial x} = \frac{\partial\psi}{\partial\theta}\frac{\partial\theta}{\partial x} + \frac{\partial\psi}{\partial r}\frac{\partial r}{\partial x} = -\frac{\partial\psi}{\partial\theta}\frac{y}{x^2+y^2} + \frac{\partial\psi}{\partial r}\frac{x}{(x^2+y^2)^{\frac{1}{2}}}$$

$$= \frac{\partial\psi}{\partial r}\cos\theta - \frac{\partial\psi}{\partial\theta}\frac{\sin\theta}{r}$$

Since

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} = \frac{\partial^2\psi}{\partial r^2} + \frac{1}{r}\frac{\partial^2\psi}{\partial\theta^2} + \frac{1}{r}\frac{\partial\psi}{\partial r}$$

Equation (3) becomes, after substitution

$$\frac{\partial^2\psi}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\theta^2} + \frac{1}{r}\frac{\partial\psi}{\partial r} - \frac{3}{\cos\theta}\left(\frac{\partial\psi}{\partial r}\cos\theta - \frac{\partial\psi}{\partial\theta}\frac{\sin\theta}{r}\right) = r^2\frac{\partial^2\psi}{\partial r^2} + \frac{\partial^2\psi}{\partial\theta^2} + 3\tan\theta\frac{\partial\psi}{\partial\theta} - 2r\frac{\partial\psi}{\partial r} = 0 \tag{4}$$

The boundary conditions for equation (4) are

$$\psi = \frac{k}{4}x^2 = \frac{k}{4}r^2\cos^2\theta \qquad \text{for} \qquad r = r_1,\ r = r_2 \tag{5}$$

Equation (5) can be handled by Bernoulli's method of particular solutions. Let

$$\psi = R(r)\cdot\Theta(\theta)$$

Then

$$\frac{\partial\psi}{\partial\theta} = R\Theta', \quad \frac{\partial^2\psi}{\partial\theta^2} = R\Theta''; \qquad \frac{\partial\psi}{\partial r} = R'\Theta, \quad \frac{\partial^2\psi}{\partial r^2} = R''\Theta$$

Substituting these latter into (4), and dividing by $R\Theta$ the following is obtained

$$\frac{r^2R''\Theta + R\Theta'' + 3R\Theta'\tan\theta - 2rR'\Theta}{R\Theta} = 0$$

or

$$r^2\frac{R''}{R} + \frac{\Theta''}{\Theta} + 3\frac{\Theta'}{\Theta}\tan\theta - 2r\frac{R'}{R} = 0$$

which breaks up into

$$r^2\frac{R''}{R} - 2r\frac{R'}{R} = -\left(\frac{\Theta''}{\Theta} + 3\frac{\Theta'}{\Theta}\tan\theta\right)$$

Let

$$r^2\frac{R''}{R} - 2r\frac{R'}{R} = m \tag{6}$$

and therefore

$$\frac{\Theta''}{\Theta} + 3\frac{\Theta'}{\Theta}\tan\theta = -m \tag{7}$$

where m is called the separation constant.

First, consider equation (6), which is Euler's homogeneous equation. Making the substitution $r = e^t$, dR/dr becomes equal to

$$\frac{1}{r}\frac{dR}{dt} \quad \text{and} \quad \frac{d^2R}{dr^2} = \frac{1}{r^2}\left(\frac{d^2R}{dt^2} - \frac{dR}{dt}\right).$$

Equation (6) becomes

$$\frac{d^2R}{dt^2} - 3\frac{dR}{dt} - mR = 0$$

which has the solution

$$\begin{aligned} R &= c_2 e^{\frac{3+\sqrt{9+4m}}{2}t} + c_1 e^{\frac{3-\sqrt{9+4m}}{2}t} \\ &= c_2 r^{\frac{3+\sqrt{9+4m}}{2}} + c_1 r^{\frac{3-\sqrt{9+4m}}{2}} \end{aligned} \tag{8}$$

where c_1 and c_2 are constants which must be determined.

Now, consider equation (7) and let $w = \sin\theta$. Then $dw = \cos\theta\, d\theta$ and $\cos^2\theta = 1 - w^2$

$$\frac{d\Theta}{d\theta} = \frac{d\Theta}{dw}\frac{dw}{d\theta} = \cos\theta\frac{d\Theta}{dw}, \quad \text{and} \quad \frac{d^2\Theta}{dw^2} = \cos^2\theta\frac{d^2\Theta}{dw^2} - \sin\theta\frac{d\Theta}{dw}$$

Equation (7) $\frac{d^2\Theta}{d\theta^2} + 3\tan\theta\frac{d\Theta}{d\theta} + m\Theta = 0$ becomes

$$\cos^2\theta\frac{d^2\Theta}{dw^2} - \sin\theta\frac{d\Theta}{dw} + 3\tan\theta\cos\theta\frac{d\Theta}{dw} + m\Theta = 0$$

or

$$(1 - w^2)\frac{d^2\Theta}{dw^2} + 2w\frac{d\Theta}{dw} + m\Theta = 0 \qquad \text{where } |w| \leq 1. \tag{9}$$

This latter equation has the self-adjoint form

$$\frac{d}{dw}\left(\frac{\Theta'}{1 - w^2}\right) + \frac{m\Theta}{(1 - w^2)^2} = 0 \tag{10}$$

Equation (9) is characterized by a regular singularity at infinity and by regular singularities at ± 1.[11] If the indicial equation of (9) is solved for the exponents at these singular points, then (9) may be written in the form of a Riemann's P-equation, as follows,

$$\Theta = P\begin{Bmatrix} 1 & \infty & -1 & \\ 0 & \alpha_1 & 0 & w \\ 2 & \alpha_2 & 2 & \end{Bmatrix} \tag{11}$$

where $\alpha_{1,2} = \dfrac{3 \mp \sqrt{9 + 4m}}{2}$

To study the integral representations of the solution of (9), introduce an Euler transformation[12] in (9) of the type

$$\Theta = \int_c K(w - t) \cdot v(t)\, dt, \quad \text{where} \quad K(w - t) = (w - t)^{\mu+1} \tag{12}$$

It is found that if (12) is a solution of (9), $v(t)$ must satisfy the equation adjoint to (9). This adjoint equation is

$$\bar{M}_t(v) = (1 - t^2)\frac{dv}{dt} - 2\mu v t = 0 \tag{13}$$

The contour c must be so chosen that the bilinear concomitant

$$[(w - t)^{\mu+1}(1 - t^2)^{-\mu+1}]_c \tag{14}$$

vanishes identically. In particular, one may choose the section of the real axis between -1 and $+1$ as the path of integration, where equation (12) equals

$$\Theta = \int_{-1}^{1} (w - t)^{\mu+1}(1 - t^2)^{-\mu}\, dt, \quad \text{and} \quad -1 \leqq \mu < 0. \tag{15}$$

However, to solve equation (9), note that it is an equation of the type[13]

$$y'' + \frac{2\nu + 1}{z^2 - 1}\, y' - \frac{n(n + 2\nu)}{z^2 - 1}\, y = 0 \tag{16}$$

[11] Modern Analysis. Whittaker and Watson, 4th ed., 1927, Cambridge University Press, pp. 197–207.

[12] Ordinary Differential Equations. E. L. Ince, 1st ed., 1927, Longmans Green and Co., p. 191.

[13] Modern Analysis, loc. cit.,[11] p. 329.

which is defined for unrestricted values of ν and n. Equation (9) has $\nu = -\frac{3}{2}$ and $m = n(n-3)$. Rewrite (9), so that it is

$$\Theta'' - \frac{2w}{w^2-1}\Theta' - \frac{n(n-3)}{w^2-1}\Theta = 0 \tag{17}$$

To obtain a particular solution of (17), try $\Theta = (w^2-1)P^{-2}_{n-2}(w)$. If, for the present, $P^{-2}_{n-2}(w)$ is denoted by P, then

$$\Theta' = 2wP + (w^2-1)P'$$

and

$$\Theta'' = 2P + 4wP' + (w^2-1)P''.$$

Substituting these values into (17), the following is obtained:

$$(w^2-1)P'' + 2wP' + \left[2 - n(n-3) - \frac{4w^2}{w^2-1}\right]P = 0 \tag{18}$$

If $4P$ is both added to and subtracted from (18), the equation

$$(1-w^2)P'' - 2wP' + \left[(n-2)(n-1) - \frac{4}{1-w^2}\right]P = 0 \tag{19}$$

is obtained. But (19) is exactly the differential equation[14] for $P^{-2}_{n-2}(w)$ and thus $(w^2-1)P^{-2}_{n-2}(w)$ is seen to be a particular solution of (17).

The associated Legendre function, $P^{-2}_{n-2}(w)$ is an algebraic function, when $n-2$ is taken as a positive integer, and the upper index, -2 is not a positive integer.[15] From the indicial equation of (9), solutions of the type $\Theta = \sum_{m=0}^{\infty} a_m w^{m+p}$ where the a_m's are constant and the m's and p's are integers, are obtained. Since, the point $w = 1$ is a regular point of the equation, then the $P^{-2}_{n-2}(w)$ where $n-2$ is a positive integer will give convergence of the solution for all a_m.

Returning to equation (8), and substituting $m = n(n-3)$,

$$R = c_1 r^n + c_2 r^{(3-n)} \tag{20}$$

is obtained. The solution ψ is the sum of the particular solutions, as follows:

$$\psi = \sum_{n=2}^{\infty} [c_1 r^n + c_2 r^{(3-n)}][(w^2-1)P^{-2}_{n-2}(w)] \tag{21}$$

[14] Modern Analysis, loc. cit.,[11] p. 324.

[15] The Theory of Spherical and Ellipsoidal Harmonics. E. W. Hobson, 1st ed., 1931, Cambridge University Press, p. 185.

From the conditions which ψ must satisfy at the boundary, when $r = r_1$ then

$$\psi = \psi_1 = \frac{k}{4} r_1^2 \cos^2 \theta$$

and when $r = r_2$,

$$\psi = \psi_2 = \frac{k}{4} r_2^2 \cos^2 \theta$$

Since the functions $\left[\frac{2n+1}{2}\frac{(n+2)!}{(n-2)!}\right]^{\frac{1}{2}} P_n^{-2}(w)$ are orthogonal, and normalized, over the interval $(-1, 1)$,[14] then

$$\psi_1 = (w^2 - 1) \sum_{n=2}^{\infty} A_n \left[\frac{2n+1}{2}\frac{(n+2)!}{(n-2)!}\right]^{\frac{1}{2}} P_n^{-2}(w) = -\frac{k}{4} r_1^2 (w^2 - 1) \quad (22)$$

and

$$\psi_2 = (w^2 - 1) \sum_{n=2}^{\infty} B_n \left[\frac{2n+1}{\cdot 2}\frac{(n+2)!}{(n-2)!}\right]^{\frac{1}{2}} P_n^{-2}(w) = -\frac{k}{4} r_2^2 (w^2 - 1) \quad (23)$$

as $\cos^2 \theta = -(w^2 - 1)$. The coefficients A_n and B_n can be found as

$$A_n = K_1 \int_{-1}^{1} P_n^{-2}(w)\, dw \quad (24)$$

and

$$B_n = K_2 \int_{-1}^{1} P_n^{-2}(w)\, dw \quad (25)$$

where

$$K_1 = -\frac{k}{4} r_1^2 \left[\frac{2n+1}{2}\frac{(n+2)!}{(n-2)!}\right]^{\frac{1}{2}}$$

and

$$K_2 = -\frac{k}{4} r_2^2 \left[\frac{2n+1}{2}\frac{(n+2)!}{(n-2)!}\right]^{\frac{1}{2}}$$

The definite integral $\int_{-1}^{1} P_n^{-2}(w)\,dw$ can be evaluated in the following manner. Make use of the formulas,[16]

$$P_n^{-m}(w) = \frac{(n-m)!}{(n+m)!} P_n^m(w)$$

and[17]

$$P_n^m(w) = (-1)^m(1-w^2)^{m/2} \frac{d^m P_n(w)}{dw^m}$$

or

$$P_n^{-2}(w) = \frac{(n+2)!}{(n-2)!}(1-w^2)\frac{d^2 P_n(w)}{dw^2} \tag{26}$$

The formula of Rodrigrue's for Legendre polynomials is[18]

$$P_n(w) = \frac{1}{2^n \cdot n!}\frac{d^n}{dw^n}(w^2-1)^n \tag{27}$$

This leads to the consideration of the integral

$$-\int_{-1}^{1}(w^2-1)\frac{d^{n+2}}{dw^{n+2}}(w^2-1)^n\,dw \tag{28}$$

Integrating (28) by parts

$$= -\left[(w^2-1)\frac{d^{n+1}}{dw^{n+1}}(w^2-1)^n\right]_{-1}^{1} + \int_{-1}^{1} 2w\frac{d^{n+1}}{dw^{n+1}}(w^2-1)^n\,dw$$

where the first term vanishes. Integrating the second term by parts again, (28) equals

$$\left[2w\frac{d^n}{dw^n}(w^2-1)^n\right]_{-1}^{1} - \int_{-1}^{1} 2\frac{d^n}{dw^n}(w^2-1)^n\,dw \tag{29}$$

Since $P_0(w) = 1$ the second expression in (29) is precisely

$$2\int_{-1}^{1} P_0(w)\cdot P_n(w)\,dw$$

[16] The Theory of Spherical and Ellipsoidal Harmonics, loc. cit.,[16] p. 99.
[17] The Theory of Spherical and Ellipsoidal Harmonics, loc. cit.,[16] p. 90.
[18] Modern Analysis, loc. cit.,[11] p. 383.

which vanishes on account of the orthogonality of the Legendre polynomials, since $n \neq 0$. Thus, the evaluation of $\int_{-1}^{1} P_n^{-2}(w)\,dw$ has been reduced to the evaluation of

$$\left[2w\frac{d^n}{dw^n}(w^2-1)^n\right]_{-1}^{1} = \left[2w\frac{d^n}{dw^n}[(w-1)^n(w+1)^n]\right]_{-1}^{1} \tag{30}$$

The rule of Leibniz for the nth derivative of a product $z = xy$ is

$$d^n z = xd^n y + ndxd^{n-1}y + \frac{n(n-1)}{1\cdot 2}d^2xd^{n-2}y + \cdots + ydx^n$$

and when this is applied to (30), the result is

$$2\{w[n!(w+1)^n + \text{terms containing both } (w-1) \text{ and } (w+1) + n!\,(w-1)^n]\}_{-1}^{1} \tag{31}$$

The center terms of (31) vanish at both limits, while the end terms equal

$$2n!\,[(2)^n + (-2)^n]. \tag{32}$$

When n is odd, expression (32) vanishes. When n is even, (32) equals

$$2^{n+2}\cdot n! \tag{33}$$

and therefore the summation of (21) will be for even integers only.

A system of functions $f_n(x)$ is said to be closed over an interval (a, b) if there is no function $\phi(x)$ not a member of the system, such that $\int_a^b f_n(x)\cdot\phi(x)\,dx = 0$ for all n. Applying this definition to the present case, it follows from (26) that the $P_n^{-2}(w)$ are simply polynomials, and since $n \to \infty$ all the even powers of w are present, and the $\dfrac{P_n^{-2}(w)}{1-w^2}$ are a closed set over the interval $(-1, 1)$.

From (24), (26), (27), and (33)

$$A_n = -kr_1^2\left(\frac{2n+1}{2}\right)^{\frac{1}{2}}\left[\frac{(n+2)!}{(n-2)!}\right]^{\frac{1}{2}} \equiv q_1 \tag{34}$$

likewise

$$B_n = -kr_2^2\left(\frac{2n+1}{2}\right)^{\frac{1}{2}}\left[\frac{(n+2)!}{(n-2)!}\right]^{\frac{1}{2}} \equiv q_2 \tag{35}$$

Then

$$c_1^{(n)}r_1^n + c_2r_1^{(3-n)} = A_n = q_1 \tag{36}$$

and

$$c_1^{(n)} r_2^n + c_2 r_2^{(3-n)} = B_n = q_2 \tag{37}$$

Solving (36) and (37) as simultaneous equations, where it is to be remembered that the n are even integers.

$$c_1^{(n)} = \frac{\begin{vmatrix} q_1 & r_1^{(3-n)} \\ q_2 & r_2^{(3-n)} \end{vmatrix}}{\begin{vmatrix} r_1^n & r_1^{(3-n)} \\ r_2^n & r_2^{(3-n)} \end{vmatrix}} \tag{38}$$

and

$$c_2^{(n)} = \frac{\begin{vmatrix} r_1^n & q_1 \\ r_2^n & q_2 \end{vmatrix}}{\begin{vmatrix} r_1^n & r_1^{(3-n)} \\ r_2^n & r_2^{(3-n)} \end{vmatrix}} \tag{39}$$

Whereupon

$$c_1^{(n)} = -\frac{k\left(\frac{2n+1}{2}\right)^{\frac{1}{2}}\left[\frac{(n+2)!}{(n-2)!}\right]^{\frac{1}{2}}\{r_1^2 r_2^{(3-n)} - r_2^2 r_1^{(3-n)}\}}{r_1^n r_2^{(3-n)} - r_2^n r_1^{(3-n)}} \tag{40}$$

and

$$c_2^{(n)} = -\frac{k\left(\frac{2n+1}{2}\right)^{\frac{1}{2}}\left[\frac{(n+2)!}{(n-2)!}\right]^{\frac{1}{2}}\{r_2^2 r_1^n - r_1^2 r_2^n\}}{r_1^n r_2^{(3-n)} - r_2^n r_1^{(3-n)}} \tag{41}$$

From (21) and (26)

$$\psi = k(1-w^2)^2 \sum_{n=2,\,4,\,6,\cdots}^{\infty} \left(\frac{2n+1}{2}\right)^{\frac{1}{2}}\left[\frac{(n+2)!}{(n-2)!}\right]^{\frac{1}{2}} \left\{\frac{[r_1^2 r_2^{(3-n)} - r_2^2 r_1^{(3-n)}]\, r^n + [r_2^2 r_1^n - r_1^2 r_2^n]\, r^{(3-n)}}{r_1^n r_2^{(3-n)} - r_2^n r_1^{(3-n)}}\right\} \frac{d^2 P_n(w)}{dw^2} \tag{42}$$

Now, study the convergence of (42) where $\psi = \psi(r, w)$. The terms

$$\frac{[r_1^2 r_2^{(3-n)} - r_2^2 r_1^{(3-n)}]\, r^n + [r_2^2 r_1^n - r_1^2 r_2^n]\, r^{(3-n)}}{r_1^n r_2^{(3-n)} - r_2^n r_1^{(3-n)}}$$

which contain r are bounded for all n. Then, the convergence of a function $f(w)$ which is represented by a series of the form $\sum_{n=2,4,6,\cdots}^{\infty} a_n P_n''(w)$ may be considered instead of (42). Since the $P_n''(w)$ are polynomials in w, $f'(w)$ is represented in the mean by $\sum_n a_n P_n'''(w)$. Then[19]

$$f(w) = \int_0^w f'(w)\, dw = \sum_{n=2,4,6,\cdots}^{\infty} a_n P_n''(w)$$

which converges uniformly. Then from (2),

$$\phi = -\frac{k}{4} r^2 w^2 + k(1-w^2)^2 \sum_{n=2,4,6,\cdots}^{\infty} \left(\frac{2n+1}{2}\right)^{\frac{1}{2}} \left[\frac{(n+2)!}{(n-2)!}\right]^{\frac{1}{2}} \left\{\frac{[r_1^2 r_2^{(3-n)} - r_2^2 r_1^{(3-n)}]\, r^n + [r_2^2 r_1^n - r_1^2 r_2^n]\, r^{(3-n)}}{r_1^n r_2^{(3-n)} - r_2^n r_1^{(3-n)}}\right\} \frac{d^2 P_n(w)}{dw^2}$$

where it is to be remembered that $w = \sin\theta$.

If the case where ϕ equals a different constant on $r = r_1$ than on $r = r_2$ is considered, then ψ_1 has the value $\frac{k}{4} x^2$ + a constant, and $\psi_2 = \frac{k}{4} x^2$ + another constant. Equations (24) and (25) will then contain an additional term involving the integral $\int_{-1}^{1} \frac{P_n^{-2}(w)}{w^2-1}\, dw$. From (28), it is seen that this leads to the consideration of $\int_{-1}^{1} \frac{d^{n+2}}{dw^{n+2}} (w^2-1)\, dw$. But

$$\int_{-1}^{1} \frac{d^{n+2}}{dw^{n+2}} (w^2-1)\, dw = \left[\frac{d^{n+1}}{dw^{n+1}} (w-1)^n (w+1)^n\right]_{-1}^{1}$$

which can also be treated by the rule of Leibniz.

In the more general case where $\psi(r, \theta)$ can have any value on the boundary, $\psi_1(r_1, \theta)$ and $\psi_2(r_2, \theta)$ the procedure would be similar, where the evaluation of the additional integral would be determined by the values of $\psi_1(r_1, \theta)$ and $\psi_2(r_2, \theta)$.

[19] The Theory of Functions. E. C. Titchmarsh, 1st ed., 1932, Oxford Press, p. 389.

Differential Equations for Electrical Engineers. P. Franklin. 1st ed., 1933, John Wiley and Son, p. 270.

GEODESICS ON POLYHEDRAL SURFACES[1]

By Philip Franklin

Introduction. The question of the existence of closed geodesics on regular surfaces homeomorphic to the sphere has been studied by Birkhoff,[2] Lusternik and Schnirelmann,[3] and Morse.[4] One of their main results is that, unless there exists an infinite family of simple closed geodesics of the same length, there are at least three distinct simple closed geodesics on such a surface. The object of this note is to study geodesics and closed geodesics on surfaces with singularities. For the most part we consider the simplest type of such surface, convex polyhedra in Euclidean three space. If we use the definition of geodesic in terms of shortest distance, we find that there always exist geodesic arcs between pairs of points. These will not pass through the vertices of the polyhedron. Simple closed geodesics do not always exist, as we show by examples. However, if we weaken the definition of a geodesic by using the property of the normal in a conventionalized sense, the theorems for regular surfaces may be generalized.

Open geodesics on convex polyhedra. Consider a convex polyhedron in ordinary space. A curve drawn on it will be called a geodesic G if for any arc AB, sufficiently small, the arc length is an absolute minimum, i.e. the arc has less length than any other arc joining A and B. We shall show that

Given any two distinct points on a polyhedron, at least one geodesic arc G can be drawn joining them. Its length will be a minimum. This will not contain any vertices of the polyhedron, except perhaps the end points of the arc.

The points of the polyhedron are either interior to a face, interior to an edge, or vertices. Those of the first two types have neighborhoods which are plane, or developable on a plane. The vertices have neighborhoods which develop on a sector, whose bounding lines correspond

[1] Presented to the American Mathematical Society.

[2] Birkhoff, Dynamical Systems, Am. Math. Soc. Coll. IX, 1927, p. 135 f.

[3] Lusternik and Schnirelmann, Paris C. R., 1929, vol. 188, p. 295, ibid., p. 534, vol. 189, p. 269.

[4] Morse, Closed Extremals, Annals of Mathematics, 1931, vol. 33, p. 549, and Morse, The Calculus of Variations in the Large, Am. Math. Soc. Coll. XVIII, p. 250 f.

to the same line on the surface. Any two points in one of the plane neighborhoods can be joined by a geodesic in the neighborhood, namely the straight line segment between them. Any two points in the neighborhood of a vertex may be joined by a class of straight line segments. If neither of the points is the vertex, one of these segments, and in particular the one of least length, will not pass through the vertex. For each point we may construct a neighborhood, N, in the form of a circle, or sector of a circle, of radius $2A$. Then form a new neighborhood, N', of radius A. By the Heine-Borel theorem we may select a finite subset of neighborhoods N' which completely cover the polyhedron. Let the least value of A for any of these be R. Then any two points at distance less than R will be in a single neighborhood N. For, the first of these points lies in one of the finite number of neighborhoods N' of radius exceeding R, so that the corresponding neighborhood N has a radius exceeding $2R$ and hence includes the second point.

Now consider any two distinct points of the polyhedron. They may be joined by various curves. As the lengths of these are all positive, they have a greatest lower bound, L, positive since it exceeds the distance between the points. Select a sequence of such curves whose lengths approach L as a limit, including no curves of length greater than $2L$. Introduce a parameter t, running from 0 to 1 for the arc, and proportional to the arc length, s. Then, for each coördinate, we have

$$(1) \qquad |x(t_2) - x(t_1)| \leqq |s(t_2) - s(t_1)| \leqq 2L\,|t_2 - t_1|.$$

Thus the sequence of functions $x(t)$ are equi-continuous, and bounded, to that we may select a convergent subset for each coördinate in turn, i.e. a sequence of curves approaching uniformly a continuous limiting curve. As this curve satisfies the Lipschitz condition (1), it has a length L', and may be approximated by an inscribed polygon of length P, where

$$(2) \qquad L' - \epsilon < P < L'.$$

Again, there are curves of the sequence whose coördinates are uniformly close to those of the limit curve and hence to the vertices of the polygon. If L_n is the length of one such curve, we may arrange matters so that

$$(3) \qquad P < L_n + \epsilon, \qquad \text{and} \qquad L_n < L + \epsilon.$$

That is, combining (2) and (3), and using the absolute minimum property of L,

$$(4) \qquad L \leqq L' < L + 3\epsilon.$$

Thus

$$L' = L,$$

and we have shown the existence of a geodesic G of minimum length joining our two points.

It follows from the property of the neighborhoods of the vertices that these geodesics will never have vertices of the polyhedron as intermediate points.

Closed geodesics G. A closed geodesic G on a polyhedron would be a closed polygon, not passing through any of the vertices of the polyhedron, and cutting the edges in such a way that the two sides cutting an edge cut it at equal angles. While there may be closed geodesics of various types on our surface, these frequently fail to exist. For example, if our polyhedron degenerates into a triangle covered twice, there is no non-self intersecting closed geodesic. For an isosceles triangle, there are closed geodesics with one self intersection; for a right triangle there are closed geodesics with two self intersections, and for a scalene triangle there are closed geodesics with three self intersections. In each case there are an infinite family of geodesics having the same length. In fact, on any polyhedron if there is one closed geodesic G of any type, an infinite number of geodesics of equal length may be drawn parallel to it and equidistant from it.

Geodesics G'. The absence of geodesics G passing through vertices suggests that we weaken our definition. We recall the normal property of a geodesic, related to the geodesic being the curve of equilibrium of a string. Since a string stretched in the form of a skew polygon under uniform tension will have the applied forces at the vertices along the bisectors of the angles, we require that these directions be "normal" to the polyhedron in a suitable sense. Along an edge of our polyhedron we regard the normal as in a sector of indetermination, in a plane perpendicular to the edge, and between the normals to the faces. Similarly, at a vertex, the normal lies in a pyramid of indetermination determined by the normals to the faces. These conventions are such that if a surface be constructed parallel to the polyhedron, made up of planes, cylinders and spheres, the "normals" at the edges of the polyhedron will agree with those of the cylinders, while those at the vertices will agree with those of the spheres for the parallel surface.

We define geodesics G' as polygons on the polyhedron whose normals at their angular points agree with the "normals" to the polyhedron as just defined. This condition necessitates that at an edge of the poly-

hedron, the two sides of the geodesic polygon G' make equal angles with the edge. Thus, any arc of a geodesic G' not containing a vertex of the polyhedron is a geodesic G. However, the curves G' may pass through vertices. In fact, if the vertices of the polyhedron are sufficiently acute to be included in a cone with vertical angle less than a right angle, the curve G' may enter and leave a vertex in any direction.

The definition of geodesics G' has one defect, in as much as it is not invariant under development. Thus, for example, the restrictions are different for the vertex of a double triangle with angle 135°, from those for one vertex of a cube, although each will develop on a sector of 270°. However, this is not too serious an objection for the surface as a whole, since a convex polyhedron can not be deformed.

Closed Geodesics G'. On a regular surface homeomorphic to the sphere, the length of the three closed geodesics may be found as follows.[3] We consider three classes of curves on the surface. The class A_1 consists of any one parameter family of closed curves which may be deformed into a class homeomorphic with the great circles on the sphere through two points. The class A_2 consists of any two parameter family of closed curves which may be deformed into a class homeomorphic with all the great circles of the sphere. The class A_3 consists of any three parameter family of closed curves which may be deformed into a class homeomorphic with all the circles on the sphere. In each of these classes A_1 there is a curve of maximum length. The greatest lower bound of these maximum lengths, for all classes A_1 gives the length c_1. Similarly we may find the lengths c_2 and c_3. For a regular surface, closed geodesics may be found having these three lengths.

Now consider our polyhedron, and construct a family of parallel surfaces approaching it as a limit. By suitable correspondence any class A_1, A_2, or A_3 on the polyhedron leads to a similar class on each of the parallel surfaces. Thus we may define numbers c_1, c_2, and c_3 for the polyhedron, which will be the limits of the corresponding numbers $c_{1'}$, $c_{2'}$, $c_{3'}$ for the parallel surfaces. In particular, if they are unequal for the polyhedron, they will, eventually, be unequal for the parallel surfaces. Now consider the curves of length $c_{1'}$. Their coördinates form a bounded and equi-continuous set, which will give a limiting curve of length c_1 on the polyhedron. This limiting curve will be a non self-intersecting closed geodesic G'. Similarly, in the case considered when c_2 and c_3 are distinct from c_1, there will be two other closed geodesics G' having these lengths. If two of these are equal, there will be either an infinite family, or three distinct geodesics on each of the parallel

surfaces; and so again, *at least three closed geodesics G' without self-intersections on any convex polyhedron.*

Let us return to the example of a polyhedron which degenerates into a triangle covered twice. If all the angles are acute, each of the three sides taken twice, and each of the three altitudes taken twice, as well as the perimeter taken once, give closed geodesics G'. For an obtuse triangle, the altitude from the obtuse angle taken twice, the corresponding side of the triangle taken twice, and the perimeter of the triangle taken once give closed geodesics G'. The side taken twice is not simple in the usual sense, but does not cross itself.

Polyhedral surfaces. The argument for the existence of geodesics G is valid for a much wider class of surfaces than polyhedrons. It applies essentially unchanged to any surface in space with only a finite number of conical points, the rest being regular. If we use some form of intrinsic coördinates in place of the Cartesian ones, we need not have the surface in three space, but it may merely be given abstractly, like a Riemann surface. However, if the vertices develop into sectors with angles greater than 2π, the arcs may contain vertices as interior points.

The discussion of geodesics G' only applies in the form given to convex surfaces in space, with a finite number of conical points.

THE GEOMETRY OF LINEAR HOMOGENEOUS PARTIAL DIFFERENTIAL EQUATIONS OF THE FIRST ORDER

By G. P. Wadsworth[1]

1. Introduction. If a Pfaffian expression in n-variables be equated to zero as

$$A_1dx_1 + A_2dx_2 + \cdots A_ndx_n = 0 \tag{1.10}$$

where $(A_1, A_2, \cdots A_n)$ are functions of the independent variables $(x_1, x_2, \cdots x_n)$, and if the variables be interpreted as the coördinates in n-space, then to each point of this space there is associated a planar element. The integral curves form a congruence and at any particular point the integral curve passing through that point is perpendicular to the planar element. If these planar elements envelope a surface the Pfaffian is integrable.

The equivalent linear homogeneous partial differential equation

$$A_1\frac{\partial f}{\partial x_1} + A_2\frac{\partial f}{\partial x_2} + \cdots A_n\frac{\partial f}{\partial x_n} = 0 \tag{1.11}$$

(where the A's are as before functions of the independent variables $(x_1, x_2, \cdots x_n)$) has for its solutions all the possible surfaces which can be passed through the integral curves. Thus we have a very much less stringent condition than in the case of the Pfaffian equation where the planar elements must line up in a special way in order to have any solution.

The geometry of the algebraic Pfaffian was studied by C. L. E. Moore and P. Franklin.[2]

The purpose of this paper is to study the solution of equation (1.11) with the particular restrictions on the coefficients $(A_1 \cdots A_n)$ that they shall be linear functions of the independent variables $(x_1 \cdots x_n)$. The canonical forms of equation (1.11) are obtained and from them the integral curves. We thus show that the integral surfaces may always be obtained if the coefficients are linear. The effect of a general affine

[1] This is in substance a doctor's dissertation, written under the direction of Professor Philip Franklin, and accepted by the department of mathematics of M. I. T., June, 1933.

[2] Journal of Mathematics and Physics, Vol. X, No. 1, 1931, p. 19.

and projective transformation upon the differential equation is studied and also the special cases where the degree of the coefficients is reduced by one, effected through a particular projective transformation. Finally, we consider the particular integral surfaces which are planes.

2. **The Three Dimensional Affine Case.** We shall show that if $n = 3$ in equation (1.11) and if the coefficients are linear functions of the three variables x, y, z, the integral surfaces may be obtained in all but one case, and in all cases if we wish to express x, y, z as functions of a parameter t.

The corresponding total differential equation to the partial differential equation (1.11) in three variables is

$$\frac{dx}{Ax + By + Cz + D} = \frac{dy}{A'x + B'y + C'z + D'} = \frac{dz}{A''x + B''y + C''z + D''} \tag{2.11}$$

If the three planes are not parallel, we may take their common point of intersection as a new origin and then

$$D = D' = D'' = 0$$

Under this condition, we can show that (2.11) is always integrable by writing it in the form

$$\frac{dx}{Ax + By + Cz} = \frac{dy}{A'x + B'y + C'z} = \frac{dz}{A''x + B''y + C''z} = \frac{dt}{1} \tag{2.12}$$

which yields the following three equations with t as an independent variable.

$$\begin{aligned} \frac{dx}{dt} &= Ax + By + Cz \\ \frac{dy}{dt} &= A'x + B'y + C'z \\ \frac{dz}{dt} &= A''x + B''y + C''z \end{aligned} \tag{2.13}$$

It is a well known fact that the general integral of this system of equations can be found from the solution of an algebraic equation and that the solution always exists.

However, we have an easy method of obtaining the actual integral curves in terms of x, y, z. The denominators of (2.12) represent a collineation in the plane and as the numerators transform in the same way as the denominators, every reduction of the collineation to standard forms also gives a reduction of the total differential equation.

The canonical forms of the collineations in the plane are well known and are determined either by considering the fixed points of the collineation, that is, whether it has one, two or three fixed points, or by means of the theory of elementary divisors. The integration of these canonical forms gives the integral curves and the solution of the partial differential equation (1.11) will be an arbitrary function of the two sets of integral curves corresponding to each form.

The integral surfaces are (including those resulting from a singular collineation)

$$
\begin{aligned}
&F[(a \ln x - b \ln y), \quad (a \ln x - c \ln y)] \\
&F\left[\left(\ln \frac{x}{y}\right), \quad (a \ln x - b \ln y)\right] \\
&F\left[\left(\frac{x}{y}\right), \quad \left(\frac{y}{z}\right)\right] \\
&F[(b \ln y - a \ln z), \quad (ax - y \ln y)] \\
&F\left[\left(\frac{y}{z}\right), \quad (ax - y \ln y)\right] \\
&F\left[(ay - z \ln z), \quad \left(\frac{a^2 x}{z} + \frac{ay - z \ln z}{z} + \frac{(\ln z)^2}{z}\right)\right] \qquad (2.14) \\
&F[(z), \quad (a \ln x - \ln y)] \\
&F\left[(z), \quad \left(\frac{ax}{y} - \ln y\right)\right] \\
&F\left[(z), \quad \left(\frac{x}{y}\right)\right] \\
&F\left[(z), \quad \left(zx - \frac{y^2}{2}\right)\right] \\
&F[(z), \quad (x - z \ln y)] \\
&F[z, y]
\end{aligned}
$$

Now suppose that in equation (2.11) all three planes are parallel, then we have the form

$$\frac{dx}{Ax+By+Cz+D} = \frac{dy}{k(Ax+By+Cz)+D'} = \frac{dz}{k''(Ax+By+Cz)+D''} \tag{2.15}$$

We may also effect a translation of the planes and reduce it to the form

$$\frac{dx}{Ax+C} = \frac{dy}{x} = dz \tag{2.16}$$

which upon integration gives sets of logarithmic curves in the different coördinate planes as the integral curves.

If, however, only two planes are parallel and the third one skew, we may take the z axis perpendicular to the two parallel planes and we arrive at the form

$$\frac{dx}{Ax+B} = \frac{dy}{x+C} = \frac{dz}{Dx+Ey+Fz} \tag{2.17}$$

Only one set of integral curves may be obtained by quadratures in this case. If, however, we place this set of ratios equal to $\frac{dt}{1}$, x, y, z may be obtained as functions of t, but the elimination of t between the three equations cannot be carried out.

Thus, in the case of a linear homogeneous partial differential equation of the first order with coefficient linear functions of the independent variables, the integral curves can always be obtained in terms x, y, z and a parameter t.

In fact, this proposition is true for all higher dimensions. For example, in four dimensions the canonical forms would be determined by the collineations in space. The integral curves will be of exactly the same general form containing as high a power of the logarithm as the cube. The forms which cannot be integrated in terms of x, y, z, will be those where we have 2 or 3 planes parallel.

3. **Projective Transformations and Homogeneous Coördinates.** The equivalent total differential equation of (1.11) in three dimensions

$$\frac{dx}{A} = \frac{dy}{B} = \frac{dz}{C} \tag{3.11}$$

where A, B, C are polynomials of the nth degree may be reduced to the form

$$\frac{x_1\,dx_4 - x_4\,dx_1}{A_1} = \frac{x_2\,dx_4 - x_4\,dx_2}{A_2} = \frac{x_3\,dx_4 - x_4\,dx_3}{A_3} \tag{3.13}$$

by the transformation

$$x = \frac{x_1}{x_4} \qquad y = \frac{x_2}{x_4} \qquad z = \frac{x_3}{x_4} \tag{3.131}$$

The polynomials A_1, A_2, A_3 are homogeneous of degree n.

If we consider the first and second equalities together, and the second and third together, we see that they are representable by the vanishing of two determinants.

$$\begin{vmatrix} dx_1 & dx_2 & dx_4 \\ x_1 & x_2 & x_4 \\ A_1 & A_2 & 0 \end{vmatrix} = 0 \qquad \begin{vmatrix} dx_2 & dx_3 & dx_4 \\ x_2 & x_3 & x_4 \\ A_2 & A_3 & 0 \end{vmatrix} = 0 \tag{3.14}$$

These two determinants suggest a more general form of which (3.14) is a special case, namely

$$\begin{vmatrix} dx_1 & dx_2 & dx_4 \\ x_1 & x_2 & x_4 \\ A_1 & A_2 & A_4 \end{vmatrix} = 0 \qquad \begin{vmatrix} dx_2 & dx_3 & dx_4 \\ x_2 & x_3 & x_4 \\ A_2 & A_3 & A_4 \end{vmatrix} = 0 \tag{3.15}$$

or generalizing still more, any third order determinant of the following matrix must vanish.

$$\left\| \begin{matrix} dx_1 & dx_2 & dx_3 & dx_4 \\ x_1 & x_2 & x_3 & x_4 \\ A_1 & A_2 & A_3 & A_4 \end{matrix} \right\| \tag{3.16}$$

If, however, (3.15) holds, then all the third order determinants of (3.16) must also vanish, subject, however, to the restriction that dx_1, dx_2, dx_3, dx_4 are independent and that no two columns of the matrix are proportional. For, if (3.15) is satisfied, we may make linear combinations of the two determinants which will result in the vanishing of the other two determinants of the matrix.

Another form of (3.15) is

$$\begin{aligned} P_1dx_1 + P_2dx_2 + P_4dx_4 &= 0 \\ Q_2dx_2 + Q_3dx_3 + Q_4dx_4 &= 0 \end{aligned} \tag{3.20}$$

where P_1, P_2, P_4, Q_2, Q_3, Q_4, are homogeneous polynomials of the nth degree with the added restriction that

$$\begin{aligned} P_1x_1 + P_2x_2 + P_4x_4 &= 0 \\ Q_2x_2 + Q_3x_3 + Q_4x_4 &= 0 \end{aligned} \tag{3.21}$$

There are also two other possible forms of (3.20) by taking the other two determinants of the matrix but they are dependent on these two.

We may check equations (3.20) by multiplying out the two determinants (3.15) and observing that we arrive at the form (3.20) with equations (3.21) satisfied.

Conversely, if we take A_1, A_2, A_3, A_4 any four homogeneous functions satisfying the relations

$$\begin{aligned} P_1A_1 + P_2A_2 + P_4A_4 &= 0 \\ Q_2A_2 + Q_3A_3 + Q_4A_4 &= 0 \end{aligned} \tag{3.22}$$

we have upon combining (3.22) (3.21) (3.20), the two forms of (3.15).

A more general expression than this can be derived from the equations

$$\begin{aligned} \sum s_i dx_i = 0 \quad \sum s_i x_i = 0 \quad \sum s_i A_i = 0 \\ \sum t_i dx_i = 0 \quad \sum t_i x_i = 0 \quad \sum t_i A_i = 0 \end{aligned} \tag{3.23}$$

where the s's and t's are homogeneous functions of degree n and the summations all run from 1 to 4. Any one of the four forms may then be derived by eliminating the corresponding dx_i, x_i, A_i, from the three sets of equations of (3.23).

If we use the identity relations of (3.22)

$$\begin{aligned} P_1(-P_2) + P_2(P_1) + P_4(0) &= 0 \\ Q_2(-Q_3) + Q_3(Q_2) + Q_4(0) &= 0 \end{aligned} \tag{3.24}$$

we obtain the original equations (2.14). This shows that in one sense all three forms are equivalent.

Let us now consider the general behavior of these different forms under projective transformations. The forms (3.11) and (3.13) are unchanged through the application of a general affine transformation but do change their form under a projective transformation. However, the two forms (3.20) and (3.21) neither change their form under a projective transformation nor under changes in the ratios of the coördinates of the form

$$(3.25)\qquad x_1 = Kx_1' \qquad x_2 = Kx_2' \qquad x_3 = Kx_3' \qquad x_4 = Kx_4'$$

where the K is any arbitrary function of the coördinates.

Since the A_i are any homogeneous functions satisfying (3.22) they are to some extent arbitrary and even if we take them of the minimum possible degree, we may replace them by $A - Lx_i$ where L is one degree lower than A_i as is seen from the form (3.15). If we begin with the coefficients of the nth degree in the form (3.20) and reduce (3.13) by the means of (3.24), we should have A_1, A_2, and A_3 of the nth degree. However, if we begin with A_1, A_2, and A_3 in (3.13) of the nth degree and go over to the form (3.20) the P_i will be in general of the degree $(n + 1)$. Thus there are two cases of partial differential equations of the nth degree, those reducible to the form (3.13) with A_i of the nth degree and those reducible to (3.20) with P_i of the nth degree. Their typical equation may be obtained by placing $dx_4 = 0$, $x_4 = 1$ in these two forms

$$(3.26)\qquad \begin{vmatrix} dx_1 & dx_2 & 0 \\ x_1 & x_2 & 1 \\ A_1 & A_2 & 0 \end{vmatrix} \equiv 0 \qquad \begin{vmatrix} dx_2 & dx_3 & 0 \\ x_2 & x_3 & 1 \\ A_2 & A_3 & 0 \end{vmatrix} \equiv 0$$

or

$$(3.27)\qquad \begin{vmatrix} dx_1 & dx_2 & 0 \\ x_1 & x_2 & 1 \\ A_1 & A_2 & A_4 \end{vmatrix} \equiv 0 \qquad \begin{vmatrix} dx_2 & dx_3 & 0 \\ x_2 & x_3 & 1 \\ A_2 & A_3 & A_4 \end{vmatrix} \equiv 0$$

For degrees higher than the first these two cases are essentially different, but for constant terms or terms of the first degree they are projectively equivalent.

Let us consider a general projective transformation

$$x_i' = \sum a_{is} x_s$$

This may be effected in the matrix as follows

$$(3.28)\qquad \begin{Vmatrix} dx_1 & dx_2 & dx_3 & dx_4 \\ x_1 & x_2 & x_3 & x_4 \\ A_1 & A_2 & A_3 & A_4 \\ 0 & 0 & 0 & 0 \end{Vmatrix} \cdot \begin{Vmatrix} a_{11} & a_{21} & a_{31} & a_{41} \\ a_{12} & a_{22} & a_{32} & a_{42} \\ a_{13} & a_{23} & a_{33} & a_{43} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{Vmatrix}$$

or carrying out the indicated operation

$$\left\| \begin{array}{cccc} \sum a_{1j}dx_j & \sum a_{2j}dx_j & \sum a_{3j}dx_j & \sum a_{4j}dx_j \\ \sum a_{1j}x_j & \sum a_{2j}x_j & \sum a_{3j}x_j & \sum a_{4j}x_j \\ \sum a_{1j}A_j & \sum a_{2j}A_j & \sum a_{3j}A_j & \sum a_{4j}A_j \end{array} \right\| \tag{3.29}$$

In this matrix all summations run from 1 to 4.

This also shows that the original matrix (3.16) is equivalent to (3.29) with

$$A'_i = \sum a_{ij}A_j \tag{3.30}$$

when the transformation is nondegenerate, that is, where the determinant $a_{ij} \neq 0$. When the coefficients are constants we may choose the a_{ij} such that

$$A'_2 = \sum a_{2j}A_j = 0, \quad A'_3 = \sum a_{3j}A_j = 0, \quad A'_4 = \sum a_{4j}A_j = 0 \tag{3.31}$$

that is

$$\left\| \begin{array}{cccc} dx'_1 & dx'_2 & dx'_3 & dx'_4 \\ x'_1 & x'_2 & x'_3 & x'_4 \\ A'_1 & 0 & 0 & 0 \end{array} \right\| \tag{3.32}$$

which after dividing them by A'_1 gives the following typical form since any three of the A's may be chosen equal to zero

$$x_4dx_2 - x_2dx_4 = 0 \tag{3.33}$$

which equations represent straight lines through the origin.

If the transformation is degenerate, that is, $a_{ij} = 0$ all four A's may be made zero and we have straight lines perpendicular to the coördinate axes.

When the coefficients are of the first degree and are not proportional we may write

$$A_i = \sum c_{ij}x_j \tag{3.34}$$

We are now able to determine a_{ij}, such that the following four relations may all be satisfied.

$$\begin{aligned} \sum a_{ij}c_{j1} &= ka_{i1} \\ \sum a_{ij}c_{j2} &= ka_{i2} \\ \sum a_{ij}c_{j3} &= ka_{i3} \\ \sum a_{ij}c_{j4} &= ka_{i4} \end{aligned} \tag{3.35}$$

by taking the k of such a value that the four equations are consistent. This requires $A_3' = kx_3'$ so that by subtracting k times the third row from the last row in each of the determinants (3.29), we have the following matrix since we may replace A_4 by zero.

$$\left\| \begin{matrix} dx_1 & dx_2 & dx_3 & dx_4 \\ x_1 & x_2 & x_3 & x_4 \\ A_1 & A_2 & A_3 & 0 \end{matrix} \right\| \tag{3.36}$$

which in non-homogeneous coördinates give us the form (3.26).

As we have classified the forms (3.26) we may now write down the projective normal forms for the A's of (3.27) which has just been shown to be projectively equivalent to (3.26) considering only the non-singular collineations. In this way we obtain the canonical forms for the A_1, A_2, A_3, A_4.

$$\begin{array}{llll} A_1 = ax & A_2 = by & A_3 = cz & A_4 = 0 \\ A_1 = ax & A_2 = by & A_3 = bz & A_4 = 0 \\ A_1 = x & A_2 = y & A_3 = z & A_4 = 0 \\ A_1 = ax + y & A_2 = ay & A_3 = bz & A_4 = 0 \\ A_1 = ax + y & A_2 = ay & A_3 = az & A_4 = 0 \\ A_1 = ax + y & A_2 = ay + z & A_3 = az & A_4 = 0. \end{array} \tag{3.37}$$

We also may obtain this form from the consideration of the two determinants (2.15). Every reduction of the collineation

$$x_i' = A_i$$

gives a similar reduction to the determinant so that following this method of attack the canonical form for the A's is the same as the canonical forms for the collineation in space. A typical form would be

$$A_1 = cx_1 \qquad A_2 = bx_2 \qquad A_3 = cx_3 \qquad A_4 = dx_4. \tag{3.38}$$

It can be readily seen that by taking certain multiples of the second row of the determinants (2.15), we are able to reduce the number of terms, so that, by appropriate combinations of rows and columns we may reduce these canonical forms to those previously found (3.27).

4. **Equations, the Coefficients of Which Are Reducible in Degree by a Projective Transformation.** In general a projective transforma-

tion will raise the degree of an equation by one. We shall now inquire what must be the restriction on the coefficients such that the degree will be lowered by one by means of a projective transformation. Let us first consider a partial differential equation in two variables with the coefficient linear functions of these variables,

$$(4.11) \qquad (Ax + By + C)\frac{\partial f}{\partial x} + (A'x + B'y + C')\frac{\partial f}{\partial y} = 0.$$

If we let $x' = \frac{1}{x}$, $y' = y/x$ and follow this with a general affine transformation we have the effect of a general projective transformation. This transforms (4.11) into

$$(4.12) \qquad \begin{aligned}&[(Ax + By)(ax + by) + K]\frac{\partial f}{\partial x}\\ &\quad + [(Ax + By)(cx + dy) + K']\frac{\partial f}{\partial y} = 0\end{aligned}$$

where K and K' consist of constant and linear terms.

If the denominator of the projective transformation is taken as the common factor $(Ax + By)$, the transformation will reduce the quadratic coefficients to ones which are linear.

Thus if we have

$$[Ax^2 + Bxy + Cy^2 + K]\frac{\partial f}{\partial x} + [A'x^2 + B'xy + C'y^2 + K']\frac{\partial f}{\partial y} = 0$$

where K and K' are again constant and linear terms, the condition that this is reducible to an equation with linear coefficients is that

$$(4.13) \qquad \begin{aligned}&Ax^2 + Bxy + Cy^2 = 0\\ &A'x^2 + B'xy + C'y^2 = 0\end{aligned}$$

shall have a common factor.

The necessary and sufficient condition that this be true is that the eliminant shall be zero, or

$$(4.14) \qquad \begin{vmatrix} A & B & C & 0\\ 0 & A & B & C\\ 0 & A' & B' & C'\\ A' & B' & C' & 0 \end{vmatrix} = 0.$$

It is a general fact that if the highest order terms of the polynomial coefficients have a common factor, regardless of the order, a special projective transformation can be found which reduces by one the degree of the coefficients.

Let us now consider the case of three variables with the coefficients quadratic in these three variables

$$\left(\sum_{i,j=1,2,3} a_{ij}x_ix_j + K\right)\frac{\partial f}{\partial x_1} + \left(\sum b_{ij}x_ix_j + K'\right)\frac{\partial f}{\partial x_2} + \left(\sum c_{ij}x_ix_j + K''\right)\frac{\partial f}{\partial x_3} = 0 \tag{4.15}$$

where again K, K', K'' are constant or linear terms, and $a_{ij} = a_{ji}$.

Suppose now that the quadratic part of the three coefficients be considered as 3 conics

$$C_1 = \sum a_{ij}x_ix_j \qquad C_2 = \sum b_{ij}x_ix_j \qquad C_3 = \sum c_{ij}x_ix_j. \tag{4.16}$$

If the three conics have a line in common, a linear combination of the three conics

$$\lambda_1C_1 + \lambda_2C_2 + \lambda_3C_3 = 0 \tag{4.17}$$

will consist of two lines. One of these lines will be common to all three conics.

Therefore, the discriminant of (4.17) will vanish for all values of the λ's. This is a necessary condition and involves the vanishing of ten invariants. Three of these invariants are the vanishing of the discriminants of the conics, which condition requires that the conics can be factored into two linear factors. Six more invariants are the three respective pairs of mutual invariants of the pairs of conics, and the last invariant is the vanishing of the cubic determinant of the cubic matrix of the 27 coefficients a_{ij}, b_{ij}, c_{ij}, where the order of the letters is immaterial.

This condition is also sufficient. For if we have for all values of the λ's

$$\lambda_1C_1 + \lambda_2C_2 + \lambda_3C_3 = (ax + by + cz)(a'x + b'y + c'z). \tag{4.18}$$

The left hand side of this equation is linear in the λ's, and therefore, the right hand side must be also. Therefore, it follows that one of the factors on the right must be independent of the λ's.

5. **Integral Surfaces Which Are Planes.** Let us first discover what are the conditions that the integral curves be straight lines. This may be effected by substitution of the general equation for the straight line

$$\begin{aligned} x &= X + p_1 t \\ y &= Y + p_2 t \\ z &= Z + p_3 t \end{aligned} \tag{5.11}$$

(where X, Y, Z are the coördinates of any point in space and p_1, p_2, p_3 are constants to be determined) into any one of the canonical forms of the equivalent total differential equation in three variables.

We find that in general there are only three planes which contain lines which are integral curves of the partial differential equation. These three planes are the coördinate planes. In special cases we have all the lines which go through the origin as integral curves.

We may verify this in a different way. The arbitrary plane

$$\begin{aligned} x &= X + p_1 + + q_1 t' \\ y &= Y + p_2 + + q_2 t' \\ z &= Z + p_3 + + q_3 t' \end{aligned} \tag{5.12}$$

when substituted in any of the canonical forms contains three straight lines which are integral curves; that is, the three lines in which the plane cuts the three coördinate planes. In special cases, if the plane goes through the origin, it may contain an infinite number of integral curves.

THE EQUIVALENT WAVE METHOD—THE USE OF VECTORS IN STUDYING ELECTRICAL TRANSIENTS

By F. Creedy[1]

It is usual to observe electrical transients by means of the ordinary or cathode-ray oscillograph. In some cases, however, such as transmission line transients, it would be desirable, were it possible, to make a continuous record of transients if a simple type of instrument, recording say on a paper roll, were available. It is not impossible to do something equivalent with the cathode-ray oscillograph but this has never yet been reduced to such a simple form that it can form part of standard equipment. It has been shown, however, by careful experimental work[2] that the form of transient usually met with can be represented sufficiently nearly by an equation of the form.

$$Ae^{at} \sin bt \quad \text{or} \quad Ae^{at} \sinh bt.$$

This fact is here made the basis of an approximate treatment of transients bearing to the accurate treatment by means of differential equations the same relation that the "equivalent sine wave" method of studying steadily periodic phenomena bears to the rigorous method using Fourier's series.

According to this method then we substitute for any surge observed by our instruments an "equivalent hyperbolic or harmonic surge." This we define to mean a surge of the above form which will give the same readings on our instruments as those actually observed.

Compare the definition of the equivalent sine waves of current and volts in any case as those waves which would yield the same ammeter, voltmeter, and wattmeter readings as we observe in that case. Such an equivalent surge then is defined by three constants A, a, and b. Theoretically, then, any three observations should be sufficient to determine it. Practically, however, unless these observations are rather carefully chosen the calculations may be rendered complex.

Assume we wish to find the "equivalent wave" to a current surge. The following observations may be taken and lead to simple calculations.

[1] University of British Columbia.

[2] L. V. Bewley, Travelling Waves in Transmission Systems, AIEE Trans., June 1931. H. L. Rorden Solution of Circuits Subject to Travelling Waves, AIEE Trans., Sept. 1932.

If a surge passes through a ballistic galvanometer, the first throw will be proportional to the integrated value M of the surge. A ballistic instrument can readily be arranged so that the surge passes through both fixed and moving coils in series when it will read the integrated square value. Or the fixed coil may be used as current coil and the moving coil as pressure coil when it will read watt-seconds. Let us read therefore by means of three ballistic instruments

M the integrated current

S the integrated square volts across the secondary of a current transformer

T the integrated current produced by the secondary of a current transformer connected in series with the instrument and a condenser.

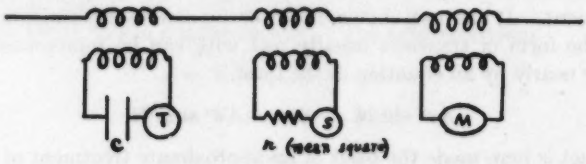


FIG. 1

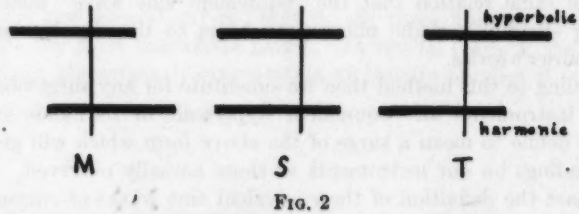


FIG. 2

Then it is shown in the Appendix that

$$a = \frac{T^2}{4S} \qquad M = \frac{-J^2Ab}{a^2 - J^2b^2}$$

$$b = \sqrt{\frac{T^4}{16S^2} + \frac{T}{M}} \quad \text{(may be imaginary)} \qquad S = \frac{A^2b^2}{4a}$$

$$A = \sqrt{\frac{T^4}{16S^2} + \frac{T}{M}}^{T} \qquad T = J^2Ab$$

For J^2 we may substitute ± 1.

This is fully explained below. S is therefore always negative since a is negative. If we take $J^2 = -1$ (harmonic surge) M is positive and T negative. If we take $J^2 = +1$ (hyperbolic surge) M is negative (unless $b > a$) and T positive.

Hence if S and T are in *the same* direction we have a harmonic or oscillatory surge. If they are in opposite directions we have a hyperbolic surge.

The three reflecting ballistic instruments might record, for instance, on a common roll of sensitized paper moved by clockwork. Preferably the damping should be fairly high so that the second swing should be noticeably less than the first. The record would then appear as in fig. 2, containing three straight lines parallel to the direction of motion of the paper with excursions at right angles whenever a surge occurs, the longer excursion measuring the first swing. A mere inspection of

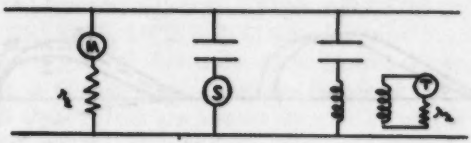


FIG. 3

the record then tells us whether the surge is harmonic or hyperbolic and also whether $b > a$ or not.

A corresponding scheme of connections to measure voltage would be that shown in fig. 3 which would employ the same formulae.

But we can connect our watt-second-meter as a wattmeter is connected and shall *not* get a reading proportional to the product of volts and amperes. In other words, there will be a function which we may call the "generalized power factor."

The simplest way to conceive how such a "generalised power-factor" can exist, when a transient wave of *EMF* strikes an impedance, is to suppose that the current and *EMF* waves start at different instants, the one being denoted by, say, $E_0\epsilon^{a_0 t} \sinh b_0 t = E$ and the other by $\epsilon^{a_0 t}(I_1 \sinh b_0 t + I_2 \cosh b_0 t) = I$. We may define an angle β by the equation $\frac{I_2}{I_1} = \tanh \beta$. This will give $I = I_0\epsilon^{a_0 t} \sinh (b_0 t + \beta)$. Methods of calculation are given below for β when I_0, a_0 and b_0 are measured as above. (See App. I.)

This conception, however, is not physically accurate since in fact both waves start together at the same instant, the current wave consisting of a wave (m) which, taken by itself, would start at another instant so that it is not zero when $t = 0$. To bring it to zero at $t = 0$ another wave (n) or "second transient" is added to it which is of such initial magnitude as to make the resultant start from zero. The resultant is shown as I fig. (4). The impressed current wave (m) will have the same constants $a_0 b_0$ as the impressed voltage wave. The constants of the wave (n) will be those of the wave which would exist if the impedance were short-circuited while containing a store of energy and allowed to return to equilibrium undisturbed. Thus if the impedance comprises a complicated circuit this wave may be of complicated form. But it is of relatively small magnitude, being only a "transient of a transient" or second transient and thus it is sufficient for practical purposes to suppose it a wave of the second order.

FIG. 4

Thus it may be desirable to measure an impressed voltage wave, a resultant current wave and the watt-seconds resulting from the two (requiring seven ballistic instruments recording on a single strip of sensitised paper) and then resolve the resultant current wave into an impressed current wave and a "second transient wave" assumed to be of the second order. How this may be done without requiring more than the above seven observations is shown in the Appendix.

Alternatively it may be desirable to calculate the current wave due to a given EMF wave from the constants of the circuit. However complex the circuit, this can easily be done for the impressed current wave by the methods given below. For the second transient wave, if the circuit is very complicated, a second order approximation must be employed, for instance by developing the impedance function in a series by Maclaurin's theorem, or otherwise, and neglecting terms above those of the second order. Since there is no occasion for very great accuracy this is easily done.

Given the constants A, a, and b, the curve may readily be plotted,

but it may be desirable to denote a little space to a graphical method of obtaining the instantaneous values, since this gives rise to a vector representation (applicable alike to harmonic and hyperbolic transients) and a method of calculation closely analogous to that used with ordinary alternating currents. In this way the use of tables is entirely dispensed with, all calculations being made by the ordinary arithmetic of complex numbers.

The graphical method described below making use of a special kind of cross section paper which the writer calls "decremental" paper enables us to apply a logarithmic decrement of any magnitude to a sine wave plotted from a vector diagram, say, without making any calculations at all.

To prepare this "decremental" paper we proceed as follows: On a piece of ordinary squared paper plot the curve $y = e^x$ (see fig. 5) and on the same sheet plot other curves whose ordinates are definite multiples of those of the original curve. These will be the curves $y = Ye^x$ where $Y = .25, .5, .75, 1.0, 2, 3, 4, 5, 6$ for instance.

A corresponding set of curves should be plotted below the x axis as shown in Fig. 5 to take care of negative values of the function to be studied. All these curves are printed once for all on the cross section paper and of course do not have to be drawn when it is desired to make a calculation.

The use of this paper may be understood as follows. Instead of the horizontal straight lines which we have for y = constant in ordinary squared paper we have substituted these curves which are denoted as "Y = constant curves."

If we draw a line (AB fig. (5)) perpendicular to the axis of x then the curve on which its extremity B rests gives the value of Y corresponding to this line. Thus as the line (retaining the same length) moves further and further to the left the value of Y associated with it becomes less and less. This is why this method is useful for studying decrements.

Now let us plot a sine wave on our decremental paper derived for instance by projecting the rotating vector V on a vertical axis. Before we can do this we have to determine what distance along OX is to be taken to represent a complete period. This depends on the decrement we desire to apply. If the distance taken to represent a complete period is great then the decrement applied will be large. If it is small only a small decrement will be applied. Thus by adjusting the time scale (which is plotted along OX) a single type of decremental paper can be used to apply a decrement of any amount. It is this, of course,

which permits us to print the curves once for all. In fig. 5 the full and dotted lines show sine curves plotted with a respective periodic time denoted by 6 or 18 units respectively giving a fairly small and a large decrement.

If Y is taken to represent the quantity under study (current I for example) then a sine wave plotted with appropriate time-scale on decre-

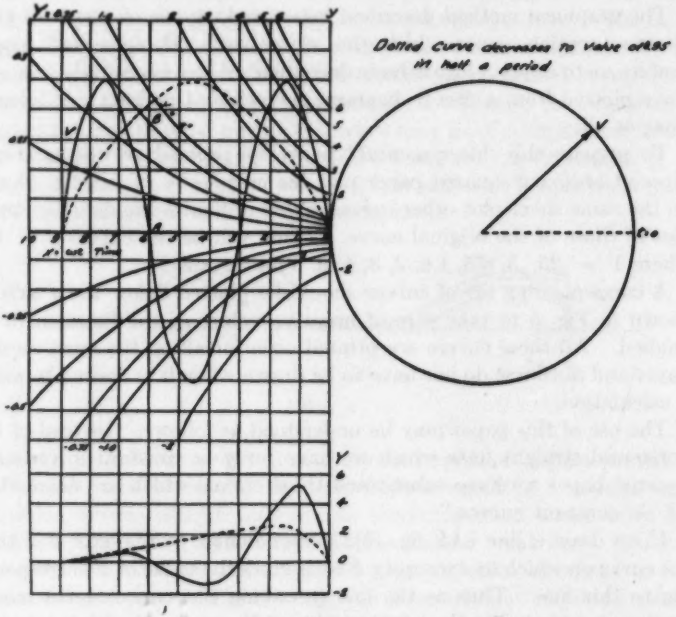


FIG. 5

mental paper is a graph of the curve with logarithmic decrement. Take the curve of fig. 5 for example.

When		Hence	
$t = 1/2$	ordinate of sine wave intersects curve	$Y = 3.$	$I = 3.$
$t = 1.25$		$Y = 5$	$I = 5$
$t = 2.25$		$Y = 3$	$I = 3$

and so on. Thus a curve which appears to the eye as a sine curve without decrement is actually a sine curve with decrement when plotted

on the special paper described. Of course it is possible to plot Y against x (or t) on ordinary squared paper when we get the ordinary decremental curve shown at the bottom of fig. (5) for both the full and the dotted curves above. But really the time required to do this is not justified as the upper curve gives all the necessary information.

It will be seen that the dotted curve, whose periodic time is denoted by 18 units, dies away almost completely in the first half period whereas the full one takes three half periods to die away. This illustrates the use of a long time scale to apply a large decrement. But if we are to

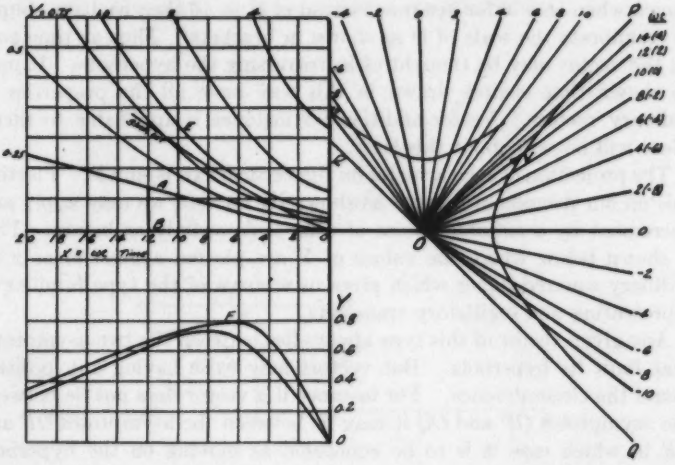


FIG. 6

be able to study transients freely we must be prepared to study the non oscillatory as well as the oscillatory type.

In the ordinary vector diagram the projection of the vector on the x axis is cos bt and that on the vertical axis sin bt. The whole vector may be denoted by

$$V(\cos bt + j \sin bt) = V e^{jbt}$$

By analogy with this let us plot a vector whose projection on the x axis is cosh bt and that on the y axis is sinh bt. The vector moves on an hyperbola whose asymptotes are lines OP and OQ at 45° to the x and y axes. Its equation in fact is $x^2 - y^2 = 1$. The distances moved

through corresponding to uniform changes in bt are also shown (Fig. 6). The values of bt may conveniently be set out along a vertical line at a convenient fixed distance from the origin. The distance from the point marked 0 to the point corresponding to any value of bt is proportional to tanh bt of course. Whenever bt has the value shown on the right, say 10, therefore, we may conceive our hyperbolic vector drawn from the origin along the line marked 10 with its extremity resting on the hyperbola. A vector N times as long would be drawn in the same direction to rest on an hyperbola which intersected the x axis at N times the distance from the origin. If the vector has the position shown when $bt = 0$ for instance instead of $bt = 10$ then we have simply to recalibrate the scale of bt as shown in brackets. Thus as time goes on the vector may be thought of as traversing the hyperbola. It may be shown that vectors drawn in this way have all the properties of ordinary vectors. Vector addition for instance is applicable to them. More will be said about this later.

The projection of such a vector on the vertical axis is sinh bt. Plotting this on our decremental paper as shown on the left, we may apply any decrement by a suitable choice of time scale exactly as before. This is shown below where the values of Y are plotted against bt or x on ordinary squared paper which gives us a curve of the type familiar as representing non oscillatory transients.

Any given vector of this type always lies between the two asymptotes that limit its hyperbola. But vectors may exist having any position round the circumference. For instance if a vector does not lie between the asymptotes OP and OQ it may lie between the asymptotes OP and OR in which case it is to be conceived as moving on the hyperbola shown $y^2 - x^2 = 1$, or a multiple of it. The positive direction for bt should now be taken clockwise instead of counter-clockwise, however. Clearly the projection of a vector to this hyperbola on the vertical would be cosh bt. Vectors opposite in sign are denoted by oppositely drawn lines in the usual way.

Thus we have a complete hyperbolic vector diagram from which we may draw our curves of instantaneous values.

Again if it were necessary to plot these hyperbolas every time it was desired to use the method, the labor would be excessive. To make it fully available therefore we require a second kind of coördinate paper in which several hyperbolas are printed once for all for all four quadrants, whose radii are equal multiples of those of the fundamental curves $x^2 - y^2 = 1$ and $y^2 - x^2 = 1$. Radial lines properly spaced will also be required of course.

These two special types of paper comprise everything that is necessary to enable transient calculations to be made as easily as steady state calculations.

The following fundamentally important proposition in circuit theory may easily be proved.

Theorem I. The Steady State vector equations to any circuit including impressed *EMFs* when not higher than the second order apply entirely unchanged in form to both oscillatory and non-oscillatory transients. The only change in passing from one of these three conditions to any other is a change in the method of calculating the individual impedance. The relations between these are invariable.

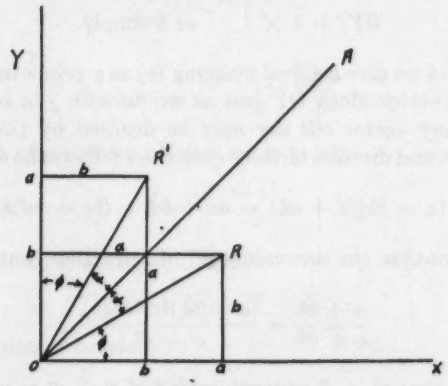


FIG. 7

It will be clear that as far as oscillatory transients go we have merely generalized the ordinary vector-diagram slightly and that it may be discussed, as heretofore, by ordinary complex numbers. It will be desirable, however, to discuss in considerably more detail how vector methods may be applied to hyperbolic transients.

Let us for the moment denote by $(a,b,)$ the vector OR having the x component a and the y component b. There will be another vector OR' often called the associate vector having the components (b, a). Clearly OR' makes the same angle with OY that OR does with OX. Or we may say that it makes the same angle α on one side of a line OA drawn at 45° to the axes that OR does on the other. We may call OR' the *reflection* of OR with respect to OA.

Let us denote by k the geometrical operation of reflection with respect to an axis at 45° to OX and OY.[3] Thus

$$k(a, b) = (b, a) \qquad k(b, a) = k \cdot k(a, b) = (a, b)$$

Hence

$$k^2 = 1 \qquad k^3 = k \qquad k^4 = 1 \qquad k^5 = k \qquad \text{etc.}$$

This is the symbolic statement of the geometrical proposition "Double reflection restores."

The reflection of a vector along OX is a vector along OY. The unit vector along OY may be symbolized by

$$OY_1 = k \times 1 \qquad \text{or } k \text{ simply.}$$

If we omit the 1 we give k a dual meaning (a) as a geometrical operation (b) as a unit vector along OY just as we do with j in other forms of geometry. Any vector OR say may be denoted by $OR = a + bk$. Multiplication and division of these quantities follows the ordinary rules.

$$(a + bk)(c + dk) = ac + bd + (bc + ad)k$$

We can "rationalize the denominator" of a fraction containing k.

$$\frac{a + bk}{c + dk} = \frac{(a + bk)(c - dk)}{c^2 - d^2}$$

We note here that $c^2 - d^2$ appears instead of $c^2 + d^2$ as with j.

$$\frac{1}{k} = \frac{k}{k^2} = k \quad \text{is another important formula.}$$

Addition of two such quantities follows the usual rule for vector additions.

$$(a + bk) + (c + dk) = (a + c) + (b + d)k.$$

In its combinations with real numbers, k by definition follows the usual commutative, associative, and distributive laws. This may be checked by considering the geometrical interpretation of the formulae.

[3] It may assist some readers if we state the equation $kZ = j\overline{Z}$ where Z is any vector and $\overline{Z}$ its conjugate.

Using the "linked brackets" notation for matrices we may state a pair of simultaneous equations in the form

$$XY = \begin{matrix}(a & b)\,(xy) \\ |c & d|\end{matrix}$$

where the rectangular array is of course a matrix. Consider some simple examples of these "binary" matrices.

We have

$$j(a + jb) = -b + ja$$

Thus in the notation of brackets

$$(-b, a) = j(a, b)$$

But we have

$$(-b, a) = \begin{matrix}(0 & -1)\,(ab) \\ |1 & 0\,|\end{matrix}$$

in the notation of matrices. Hence we may put

$$j = \begin{matrix}(0 & -1) \\ |1 & 0\,|\end{matrix}$$

thus proving that j is a matrix. Again

$$k(a + kb) = b + ka$$

or in the notation of brackets

$$(b, a) = k(a, b).$$

In matricular form we have

$$(b, a) = \begin{matrix}(0 & 1)(ab) \\ |1 & 0|\end{matrix}$$

so that

$$k = \begin{matrix}(0 & 1) \\ |1 & 0|\end{matrix}$$

The operation of "taking the Conjugate" may be similarly treated. Denote it temporarily by m so that

$$a - jb = m(a + jb) \qquad m(a - jb) = m^2(a + jb) = a + jb$$

$$\text{or } m^2 = 1.$$

In the notation of brackets

$$(a, -b) = m(a, b)$$

or

$$(a, -b) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (ab)$$

so that

$$m = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Again 1 is clearly the identical matrix.

$$(a, b) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} (ab)$$

so that

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$k(a, b) = (b, a) \qquad j(b, a) = (-a, b) = jk(a, b)$$

$$j(a, b) = (-b, a) \qquad k(-b, a) = (a, -b) = kj(a, b).$$

So that $(a, -b) = kj(a, b) = -jk(a, b) = m(a, b)$. (Because $(-a, b) = -(a, -b)$.)

Thus we have $kj = -jk = m$ and $m^2 = k^2 = 1 \quad j^2 = -1$ from which may readily be deduced the complete multiplication table of the system which is clearly a "skew commutative" one.

Any expression $\phi = x + jt + ky + mv$ may be written

$$\phi = x + jt + k(y + jv) \text{ since } kj = m.$$

Let $[\phi^2]_1$ be the "numerical" part of ϕ^2 i.e. that containing neither k j, or m.
Then

$$[\phi^2]_1 = x^2 + y^2 + z^2 - t^2$$

But this is the expression for "interval" in relativity theory and thus the possibility is opened out of illustrating the four dimensional relativity manifold by constructions in the plane. While this may be developed elsewhere, it is far from the purpose of the present paper.

It is desirable to emphasize that the matrix j or quadrantal versor is not identical with the $\sqrt{-1}$ of ordinary algebra.[4] In a two dimensional vector system using j, we can consider the points at which two curves intersect and determine them in the form $a_r + b_r j$. But if they fail to intersect a_r and b_r become imaginary, and the $\sqrt{-1}$ thus introduced must not be identified with j. If this is done these imaginary points would become real points in the plane, which is not what the equations mean. These real points have significance of course, but it is of another kind. We prefer therefore to regard $\sqrt{-1}$ as essentially uninterpretable geometrically and to use it merely to complete the system of common algebra. Notwithstanding, therefore, that

$$kj = -jk$$

we shall write

$$k\sqrt{-1} = \sqrt{-1}\,k$$

there being no objection to this since we assign no geometrical but only an algebraic meaning to such an expression. It can of course be written

$$\sqrt{-1}\,k = \begin{pmatrix} 0 & \sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}$$

or

$$\sqrt{-1}\,j = \begin{pmatrix} 0 & -\sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}$$ which is *not* equal to -1.

$\sqrt{-1}$ is not a matrix.

Consider the function e^{kx}. Expanding this we get

$$e^{kx} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots + k\left(x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots\right)$$

$e^{kx} = \cosh x + k \sinh x$. (Compare $e^{jx} = \cos x + j \sin x$.) This at once yields the hyperbolic transformation formulae. Let the "magnitude" of a vector (a, b) be defined as $\sqrt{(a^2 - b^2)}$. This is clearly quite different from its length.

e^{kx} therefore represents a vector whose components are $\cosh x$ and $\sinh x$ so that it makes a "hyperangle" x with OX. Since $\cosh^2 x -$

[4] This distinction was apparently first made by Sir W. R. Hamilton in his "biquaternions."

$\sinh^2 x = 1$, it has a constant "magnitude" in the sense defined above. It is in fact the vector to the unit hyperbola.

Any vector may with certain limitations be written in the polar form $Re^{k\theta}$, R being its "magnitude."

The transformation from Cartesian to polar coördinates proceeds as follows:

Any vector may be written

$$a + kb = \sqrt{(a^2 - b^2)}\left(\frac{a}{\sqrt{(a^2 - b^2)}} + \frac{kb}{\sqrt{(a^2 - b^2)}}\right)$$

We can find always an angle θ such that

$$\cosh\theta = \frac{a}{\sqrt{(a^2 - b^2)}}; \qquad \sinh\theta = \frac{b}{\sqrt{(a^2 - b^2)}}; \qquad \tanh\theta = \frac{b}{a}$$

The limitation just referred to is due to the fact that if $b > a$, the expression $\sqrt{(a^2 - b^2)}$ will be imaginary. This limitation may be overcome as follows.

Let $b > a$. Then $a + kb = k(b + ka)$. This may be written

$$k\sqrt{b^2 - a^2}\left(\frac{b}{\sqrt{(b^2 - a^2)}} + k\frac{a}{\sqrt{(b^2 - a^2)}}\right)$$

or

$$k\,R\,e^{k\phi}$$

if $\cosh\phi = \dfrac{b}{\sqrt{(b^2 - a^2)}}$ $\sinh\phi = \dfrac{a}{\sqrt{(b^2 - a^2)}}$ these being by hypothesis real quantities.

Now $k\,R\,e^{k\phi}$ Fig. 8 is the "reflection" of the vector $R\,e^{k\phi}$ with respect to the asymptote OA. That is as ϕ varies from 0 to $+\infty$ $R\,e^{k\phi}$ will move along the hyperbola H_1 from the point K till it coincides with the asymptote OA when $\phi = +\infty$. Simultaneously $k\,R\,e^{k\phi}$ will move from the point M when $\phi = 0$ to coincide with the asymptote OA when $\phi = +\infty$ always having such a position that $ROA = R'OA$.

If we choose to consider the hyperbolic angles from OC to OA as real then on the same basis those from OA to OD must be regarded as imaginary and so on.

This illustrates, as pointed out above, the fact that appears in many parts of analysis, that $\sqrt{-1}$ cannot *always* be interpreted by the quadrantal versor j but must sometimes be left as merely the uninterpreted

imaginary. It merely means that vectors in the quadrants *AOD* and *BOC* are somehow different in quality from those in *COA* and *BOD* though no less useful in practice.

Since in all these investigations we take hyperbolic angle proportional to time, any vector lies along *OC* when $t = -\infty$ and along *OA* when $t = +\infty$ having traversed the hyperbola in the meantime. Hence if it is in the quadrant *COA* at all there is no instant at which it can leave it. Similarly if it is in the quadrant *AOD*, it can never leave it. Thus the real quadric consisting of the two asymptotes *AB* and *CD* divides space (the plane) into two distinct regions having no communication with each other. These are often called space and antispace and

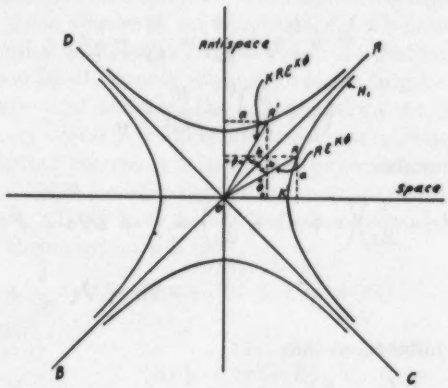


FIG. 8

vectors in them are called vectors and antivectors. The facts stated in the last paragraph are of course characteristic of any system of hyperbolic non-euclidean geometry, a simple case of which we are studying.

They have enabled us to understand clearly what is meant by the vector $Re^{k\theta}$ when R and θ are imaginary. This simply means that the vector is in antispace. Its Cartesian coördinates of course continue to be real.

The conception of "magnitude" may be still further clarified as follows:

We have seen that the vector $\cosh \theta + k \sinh \theta$ traces out the unit hyperbola as θ varies *all the radii of this hyperbola having unit magnitude* notwithstanding the very great variations in length. We may alter-

natively define the "magnitude" of any vector as *the ratio of its length to the (variable) length of the unit magnitude vector having the same direction.*

It will be desirable at this stage to take a still further step in generalization and introduce the further symbol

$$J = j \text{ or } k \qquad \text{so that} \qquad J^2 = \pm 1$$

It will be found that all vector relations can be worked out in terms of J and are then susceptible of two interpretations, a hyperbolic and a trigonometric one.

The Quadratic in $q = r + Js$. Let

$$aq^2 + bq + c = 0 \qquad (a, b, c \text{ real})$$

$$q^2 + \frac{b}{a}q + \frac{b^2}{4a^2} = \frac{b^2}{4a^2} - \frac{c}{a}$$

$$\left(q + \frac{b}{2a}\right)^2 = \frac{b^2}{4a^2} - \frac{c}{a}$$

that is a real number.

$$\left(r + Js + \frac{b}{2a}\right)\left(r + Js + \frac{b}{2a}\right) = r^2 + 2rJs + J^2s^2 + 2(r + Js)\frac{b}{2a} + \left(\frac{b}{2a}\right)^2$$

This is a real number, so that

$$2Js\left(r + \frac{b}{2a}\right) = 0. \qquad r = -\frac{b}{2a}$$

Substituting $r = -\frac{b}{2a}$ and cancelling

$$J^2s^2 = \frac{b^2}{4a^2} - \frac{c}{a}$$

so that

$$q = r + Js = -\frac{b}{2a} \pm J\sqrt{\left(\frac{b^2}{(2a)^2} - \frac{c}{a}\right)\frac{1}{J^2}}$$

It may be asked: Are there not other roots to this equation corresponding to $s = 0$, i.e. the ordinary roots of elementary algebra? The answer to this is that we have excluded these roots by our definition in which we took $q = r + Js$. Of course, if we do not exclude any

roots by definition and regard q as the general binary matrix, there are an infinite number of roots, the equation being that of a locus. This question has been much studied elsewhere and it is unnecessary to go further into it here.

If we put

$$q = Re^{J\theta} = \frac{-b \pm J\sqrt{(b^2 - 4ac)\,\frac{1}{J^2}}}{2a}$$

$$R^2 = \frac{c}{a}$$

R may be either the "magnitude" or the length of q according to whether $J = k$ or j. Since we wish to work in terms of J it is better to generalize the notation usual with complex numbers to cover both cases. We may therefore conveniently denote the magnitude or length of q by $| q |$. If $q = r + sJ$ let $r - sJ$ be denoted by $\bar{q}$ and called the conjugate of q. We shall find J very useful in working numerical examples since often we do not know whether the result is hyperbolic or oscillatory till the work is finished. But when seeking definite geometrical interpretations it is better to use k. If we have two hyperbolic vectors $A = a + bk$ and $C = c + dk$ their product will be

$$AC = (ac + bd) + k(bc + ad)$$

and its real part

$$[AC]_1 = ac + bd.$$

This may easily be proved to be equal to the product of their lengths by the cosine of the angle between them—the well known "scalar product" of vector algebra. The solution of any differential equation of the second order may be written (if real)

$$e^{at}(C \cosh bt + D \sinh bt)$$

or in our notation

$$[e^{at}(C + kD)e^{kbt}]_1 = [A\,e^{qt}]_1$$

if

$$A = C + kD$$

$$q = a + kb$$

Hence if in an electrical problem we limit ourselves to equations of the second order, the current must be expressible as above.

Since in any circuit the steady state impressed *EMF* is always balanced by the steady state voltages due to the steady state currents there is seldom any impressed *EMF* producing the transients. Thus the various *EMFs* in every circuit due to these currents must sum to zero or the total impedance of the circuit must be zero.

Since we suppose that the current can always be written Ae^{qt} the general differential equation with which we have to deal may be written

$$[(z_1 + z_2 + z_3 + \cdots z_n)\, Ae^{qt}]_1 = 0.$$

The same fact referred to above may be alternatively derived from this equation.

The current vector Ae^{qt} will have every possible direction in "space" or "antispace" (not passing from one to the other) for some value of t between $+\infty$ and $-\infty$. Nor is it ever zero. Hence if the expression in brackets above is to be zero at all instants we must have

$$z_1 + z_2 + z_3 + \cdots z_n = 0.$$

Thus in the analysis of transients an analogue of Kirchhoff's law is the statement *The sum of the transient impedances round any circuit must be zero.*

Of course Kirchhoff's laws continue to hold but the above is a convenient restatement of them for our present purpose.

For instance if we multiply each of the impedances $z_1 z_2$ etc. by the current Ae^{pt} we get the voltage across it and these voltages sum up to zero according to Kirchhoff's law.

Thus the equation $z_1 + z_2 + z_3 \cdots + z_n = 0$ may be regarded as a summation of voltages *at the instant* $t = 0$ *when the current is unity.*

DETERMINING THE INSTANTANEOUS VALUES

Suppose we have an expression $[Ae^{bkt}]_1 = 0$. e^{bkt} is a vector traversing the unit hyperbola and thus having a variable length while A is our extranormal arbitrary constant. Let $\mathcal{A} = qA$, $\mathcal{A}$ being the vector, A its length and q the unit vector having the same direction. Then clearly Ae^{kt} will traverse a hyperbola of semi-diameter A. By the above definition $[q(Ae^{bkt})]_1$ will be the product of the length of q (which is unity) on the projection of Ae^{bkt} on the axis of q.[5] If for e^{bkt} we substitute $e^{(a+bk)t} = e^{qt}$ we clearly only multiply by a numerical factor which is a function

[5] As an illustration of this note that $[ke^{ks}]_1 = \sinh x$, that is the projection of e^{ks} on the k axis which accords with the rule.

of t and do nothing which requires a change in the rule. Now consider an expression such as

$$[(z_1 + z_2 + z_3 + \cdots)Ae^{pt}]_1 = 0$$

which we saw represented the complete solution of our most general differential equation.

Suppose the terminal conditions are

(1) When $t = 0$ the projection of z_1 on the axis of q is to be C_1

(2) When $t = 0$ the projection of z_2 on the axis of q is to be C_2.

This gives us (1) $[z_1 q]_1 = C_1$; (2) $[z_2 q]_1 = C_2$ since $e^{pt} = 1$ when $t = 0$.

Although it is rather tedious to solve for q analytically the geometrical solution is exceedingly simple.

Choose such a direction for A that the projections of z_1 and z_2 on it have a ratio $\frac{C_1}{C_2}$ and such a scale for the diagram (i.e. such a length for A) that,

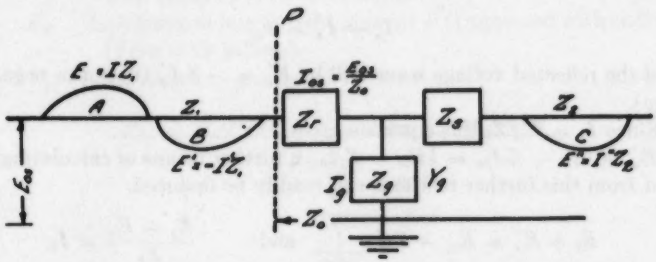


FIG. 9

to this scale, the projections have the actual magnitudes desired viz. C_1 and C_2 respectively.

Thus the graphical method of determining the arbitrary constants is very simple and particularly so when C_1 or C_2 is zero when of course A must be perpendicular to z_1 or z_2. For values of t other than zero projection on the axis so chosen will give the instantaneous values of the function. Tensor methods permit generalisation to n $(n \to \infty)$ dimensions.

SOME TRANSMISSION LINE PROBLEMS

If a travelling wave A of voltage E and current I (shown in fig. (9) as non-oscillatory) strikes an impedance inserted in a transmission line shown for generality as a T circuit then a part of it B is reflected and the rest C transmitted.

If the impedances are as shown Z_0 being the impedance of the whole circuit to the right of the point P, Z_1 being the surge impedance of the circuit to the left of P (a numeric equivalent to an ohmic resistance) and Z_2 that of the circuit to the right of the impedance then all the impressed voltages and currents to the right of P may be determined by the equations used in steady state analysis on the assumption that a voltage

$$E_{0s} = 2E_s$$

is applied at the line terminal producing a current

$$I_{0s} = \frac{E_{0s}}{Z_0 + Z_1}$$

In Z_r the reflected wave of current I' may be determined from the equation of continuity (Kirchhoff's law).

$$I_s + I'_s = I_{0s}$$

and the reflected voltage wave will be $E'_s = -Z_1 I'_s$ (note the negative sign).

Since $I_s = E_s/Z_1$ this equation gives

$E'_s = E_s - Z_1 I_{0s} = \frac{1}{2}E_{0s} - Z_1 I_{0s}$, a further means of calculating E'_s and from this further relations can readily be deduced.

$$E_s + E'_s = E_{0s} - Z_1 I_{0s} \qquad \text{and} \qquad \frac{E_s + E'_s}{Z_0} = I_{0s}$$

i.e. the sum of the incident and reflected voltages divided by the impedance to the right of P gives the inflowing current. Both these relations are sometimes useful. We have also

$$I_{0s} = (E_s - I_{0s}(Z_1 + Z_r))\, Y_0 = I_{0s}\,(Z_0 + Z_r)\, Y_0$$

and

$I_{0s} = I_{0t} + I''_s$ with $E''_s = Z_2 I''_s$ to complete the analysis by determining I''_s and E''_s.

Many of these relationships are shown in Fig. 9. A full list of symbols is as follows:

E_s Incident wave of voltage (Impressed with suffix s) (Free with suffix t).

E'_s E'_t Reflected wave of voltage (Impressed with suffix s) (Free with suffix t).

E''_s	E''_t	Transmitted wave of voltage (Impressed with suffix s) (Free with suffix t).
I_s		Incident wave of current (Impressed with suffix s) (Free with suffix t).
I'_s	I'_t	Reflected (Impressed with suffix s) (Free with suffix t).
I''_s	I''_t	Transmitted (Impressed with suffix s) (Free with suffix t).
I_{gs}	I_{gt}	Current to ground (Impressed with suffix s) (Free with suffix t).
Z_1		Surge impedance of left hand portion of line.
Z_2		Surge impedance of right hand portion of line.
Z_s	Z_{st}	Impedance of left hand arm of T circuit (Impressed with suffix s) (Free with suffix t).
Z_s	Z_{st}	Impedance of right hand arm of T circuit (Impressed with suffix s) (Free with suffix t).
Z_g	Z_{gt}	Impedance of grounded hand arm of T circuit (Impressed with suffix s) (Free with suffix t).
Z_0	Z_{0t}	Impedance of line to right of point P (Impressed with suffix s) (Free with suffix t).
Y_{gs}	Y_{gt}	Admittance to ground of T circuit.

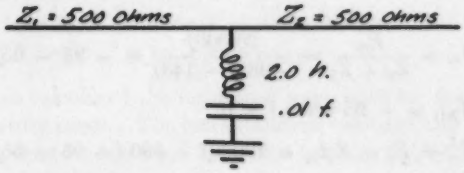


FIG. 10

It will be seen that the quantities referring to the impressed wave are distinguished from those relating to the free wave by the suffixes s and t respectively. By way of a practical example let us take such a circuit as is shown in fig. 10. Here $Z_r = Z_s = 0$ and the ground circuit contains inductance L and capacity C in series.

Let $L = 2 \quad C = .01 \quad Z_1 = Z_2 = 500.$

Let $Ee^{gt} = 10000\, J\epsilon^{(-.52 + .48J)t}$ be the impressed wave.

In the calculations we shall substitute $J^2 = \pm 1$ and then complete the oscillatory and non-oscillatory cases (where the negative or positive sign is taken) separately.

$$(-.52 + 48J)t$$

$$Z_0 = \frac{1}{\dfrac{1}{Lq + \dfrac{1}{Cq}} + \dfrac{1}{Z_2}} = \frac{(LC_q^2 + 1)\, Z_2}{Z_2 Cq + LCq^2 + 1}$$

$$q = -.52 + .48J \qquad q^2 = (-.52 + .48J)(-.52 + .48J)$$
$$= .27 \pm .23 - .5J.$$

Oscillatory Case	Non Oscillatory Case
$q^2 = .04 - .5J$	$q^2 = .5 - .5J$

Confining ourselves to the Oscillatory case, $J = j$.

$$Z_0 = \frac{(02(04 - .5j) + 1) \times 500}{3(-.52 + .48j) + 02(04 - .5j) + 1}$$
$$= \frac{500\ .4 - 5j}{-1.6 + 2.4j}$$
$$= -98.0 - 143j.$$

Observe owing to the decreasing wave the real part of Z_0 can be negative.

$$I_{0_e} = \frac{E_{0_e}}{Z_0 + Z_1} = \frac{20.000j}{-98.0 - 143j} = -95 - 65j$$

$[I_{0_e}]_1 = -95$ when $t = 0$.

$$E'_e = E_e - Z_1 I_{0_e} = 10000j - 500\,(-95 - 65j)$$
$$= 47500 + 42500j.$$

$$I'_e = \frac{E'_e}{Z_1} = \frac{-(47500 + 42500\,j)}{500} = -95 - 85j$$

$$I_e = \frac{E}{Z_1} = \frac{10000\,j}{500} = 20\,j \quad \text{so we have}$$

$$I_e + I'_e = I_{0_e} \text{ as a check.}$$

$$I''_e = \frac{E_e + E'_e}{Z_1} = \frac{47500 + 52500\,j}{500} = 95 + 105j.$$

Since Z_r and Z_s are zero in this case, $E_e + E'_e = E''_e$.

$$I_{g_e} = I_{0_e} - I''_e = -190 - 170\,J.$$

$[E_e + E'_e]_1 = 47500$ when $t = 0$.

Non-Oscillatory Case ($J = k$)

$$Z_0 = \frac{(1.01 - .01k)\,500}{500(-.0052 + 0048k)} = \frac{505 + 5k}{-1.59 + 2.39k} = 248 + 376k.$$

$$I_{0s} = \frac{E_{0s}}{Z_0 + Z_1} = \frac{20{,}000k}{748 + 376k} = -18 + 35.8k$$

$[I_{0s}]_1 = -18$ when $t = 0$.

$$E'_s = E_s - Z_1 I_{0s} = 10{,}000k - 500\,(-18 + 35.8k)$$
$$= 9000 - 7900k.$$

$$I'_s = -\frac{E'_s}{Z_1} = -\frac{9000 - 7900k}{500} = -18 + 15.8k.$$

$I_s = \dfrac{E_s}{Z_1} = \dfrac{10{,}000k}{500} = 20k$ so we have $I_s + I'_s = I_{0s}$ as a check.

$$I_s 2 = \frac{E_s + E'_s}{Z_1} = \frac{9000 + 2100k}{500} = 18 + 4.2k$$

Since Z_r and Z_s are zero in this case $E_s + E'_s = E''_s$

$$I_{gs} = I_{0s} - I''_s = -36 - 31.6\,k.$$

$[E_s + E'_s]_1 = 9000$ when $t = 0$.

Thus we have calculated the impressed wave both for the oscillatory and non-oscillatory cases. The instantaneous values of any such wave, as pointed out, may be expressed by such a formula as $[Ee^{qt}]_1$ which has the value $[E]_1$ when $t = 0$.

We have examined the values of $[I_{0s}]_1$ and $[E_s + E'_s]_1$ which correspond to the instant $t = 0$ and found they were not zero. But the physics of the problem imperatively demand that at $t = 0$ which is the moment when the wave strikes the impedance we must have the currents, voltages, etc. in it all equal to zero.

In order that this condition may be satisfied another wave system is necessary—the Free wave system as it may be called. The free wave system is the oscillation or surge which would exist if with the current and voltage at the transition point $I'_t E'_t$ having values equal to the negatives of those given by the impressed waves $[I_{0s}]_1 [E + E'_s]_1$ when $t = 0$ the line were shortcircuited at both ends and the current allowed to die away. Clearly the free wave will have the same form whether the impressed wave is oscillatory or not.

Since we have assumed $Z_1 = Z_2$ we have a surge impedance of 250 ohms for both in parallel.

The voltages round this paralleled circuit will be

$$250 \cdot I_{ot} + 2q' I_{ot} + \frac{1}{.01} q' I_{ot} = 0$$

$$(2q'^2 + 250q' + 100) I_{ot} = 0.$$

This equation can only be satisfied by the particular value of q' found by solving this quadratic.

This gives $q' = \dfrac{-250 \pm k\sqrt{250^2 - 800}}{4} = 62.5 + 62k.$

The terminal voltage $E'_t = E''_t$ will be $I_{ot} Z_{ot} = I_{ot}\left(Lq' + \dfrac{1}{Cq'}\right)$

$$= 2(-62.5 + 62k) + \frac{1}{-.625 + .62k} I_{ot}$$

$$= -250 \times I_{ot} \text{ on short circuit conditions.}$$

Under the conditions considered in the present example

$$Z_{0t} = \frac{1}{\dfrac{1}{Z_{ot}} + \dfrac{1}{Z_1}} = \frac{Z_2 Z_{ot}}{Z_2 + Z_{ot}} = -500$$

$E_t = I_t = 0$ since there can be no "incident free wave."

If we write $I'_t = (i_{1t} + k i_{2t})$

$E'_t = (e_{1t} + k e_{2t})$

the equation $I'_t = Z_0 E'_t$ can be written

$$-500\ (i_{1t} + k i_{2t}) = e_{1t} + k e_{2t}.$$

From this point on we have two calculations to make, one for the oscillatory and the other for the non-oscillatory case.

Oscillatory case.

Here we had $[I_{0e}]_1 = -95$ $\quad [E + E'_e]_1 = 47500$

Therefore $i_1 t = 95$ $\quad e_{1t} = -47500.$

i_{2t} and e_{2t} may be arbitrary or zero.

Separating the above equation into parts, one with and the other without the factor k, gives

$$-500\, i_{11} = -47500$$

$$-500 \times 95 = e_{11} = -47500 \text{ which checks.}$$

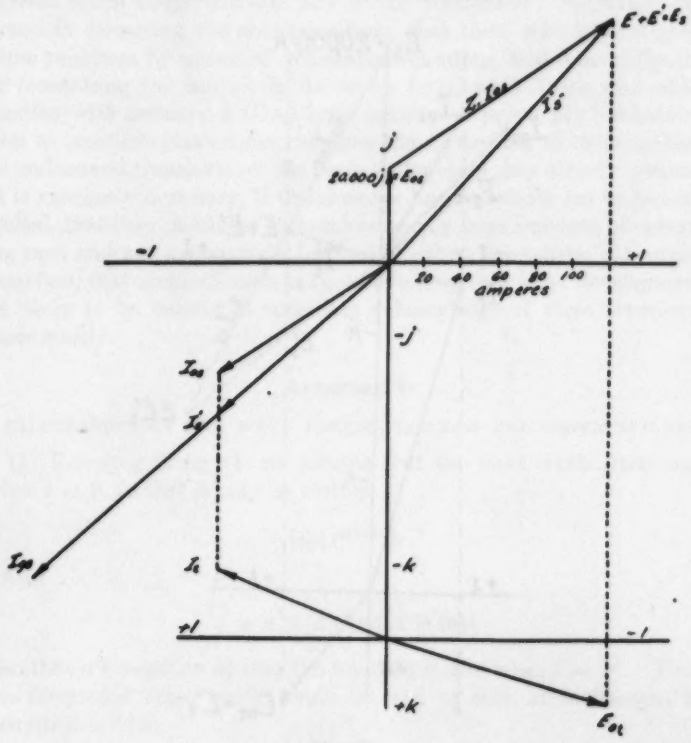


FIG. 11

Non-Oscillatory Case.

In this case we had

$$[I_{00}]_1 = -18 \qquad [E + E'_0]_1 = 9000.$$

Therefore $i_{11} = 18$ $\qquad e_{11} = -9000.$

This gives

$$-500\, i_{1t} = -9000$$

$$-500 \times 18 = e_{1t} = -9000 \text{ which checks again.}$$

The two vector diagrams fig. 11 for the oscillatory and fig. 12 for the non-oscillatory case are shown herewith. In both of these the upper

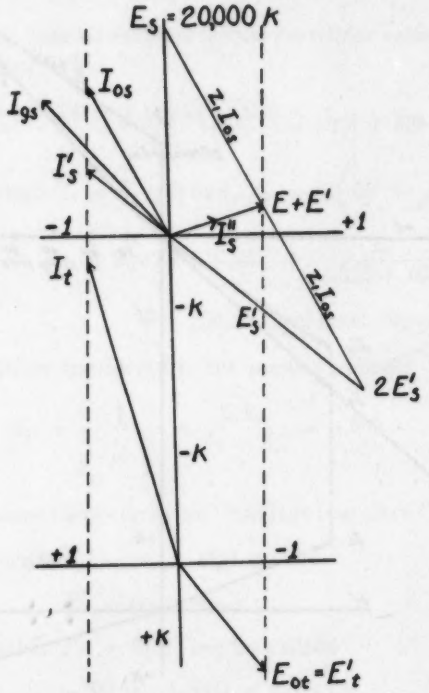


FIG. 12

part shows the quantities relating to the impressed wave and the lower those relating to the free wave. In order to show corresponding quantities directly beneath each other the positive direction for the x axis has been taken to the left and that for the y axis downward, in the lower figure instead of the more usual directions.

These vector diagrams may be changed into curves showing instan-

taneous values by the graphical methods described above which are very concise and simple to apply when the two required kinds of coördinate paper are available.

A direct comparison of the two, however, is very instructive. The current and voltage scales for the oscillatory wave are one half those for the non-oscillatory case. Owing to the more rapid variations of the current much larger currents flow in the condenser. The writer has found in discussing the above methods, that those who have studied these problems by means of differential equations, find some difficulty in recognizing the subject in its vector form, while those who while familiar with ordinary A.C. analysis, have never previously paid attention to transient phenomena, welcome these processes as helping them to understand transients on the basis of the ideas they already possess. It is extremely necessary, if transmission line transients are to be controlled, that they should be well understood by large numbers of operating men and not merely by a few highly expert specialists. It seems, therefore, that methods such as the above deserve a fuller development as likely to be helpful in spreading a knowledge of these transients more widely.

Appendix I

Calculation of the wave constants from the observations

(1) Referring to fig. (1) we assume that the wave starts from zero when $t = 0$, so that it may be written

$$[AJ\epsilon^{(a+Jb)t}]_1$$

where

$$q = a + Jb \text{ and } A \text{ is real}$$

also that a is negative so that the function is zero when $t = \infty$. Then the integrated value which would be read on such an instrument as described will be

$$M = \int_{\infty}^{0} [AJ\epsilon^{qt}]_1 = \left[\left[\frac{AJ}{q}\epsilon^{qt}\right]_1\right]_{\infty}^{0} = \left[\frac{AJ}{q}\right]_1 = \frac{-J^2Ab}{a^2 - J^2b^2}$$

(2) The voltage across the secondary of the current transformer will be

$$E = [AJq\epsilon^{qt}]_1$$

If the instrument is connected in series with a condenser of capacity C so that $CE = Q$. The instrument reads

$$\int_{\infty}^{0} \iota e dt = [Q]_{1\infty}^{0} = [CE]_{1\infty}^{0}$$

Thus the integrated value will be $T = [AJq]_1 = AJ^2b$ ignoring calibration constant C.

(3) Integrated Square Induced volts S.

$$\text{Induced volts} = [AJqe^{qt}]_1 = \tfrac{1}{2}A(Jqe^{qt} + \bar{J}\bar{q}e^{\bar{q}t})$$

$$\text{Volts squared} = \frac{A^2}{4}(J^2q^2e^{2qt} + \bar{J}^2\bar{q}^2e^{2\bar{q}t} + 2J\bar{J}q\bar{q}e^{(q+\bar{q})t}).$$

$$S = \text{Integrated Value} = \left(J^2\frac{q^2}{2q} + \frac{\bar{J}^2\bar{q}^2}{2\bar{q}} + 2J\bar{J}\frac{q\bar{q}}{q+\bar{q}}\right)\frac{A^2}{4}$$

$$= J^2\frac{A^2}{4}\left(\frac{1}{2}(q+\bar{q}) - \frac{2q\bar{q}}{q+\bar{q}}\right)$$

$$= \frac{J^2A^2}{8}\frac{(q+\bar{q})^2 - 4q\bar{q}}{q+\bar{q}}$$

$$= \frac{J^2A^2}{8}\frac{q^2 + \bar{q}^2 + 2q\bar{q} - 4q\bar{q}}{q+\bar{q}}$$

$$\frac{J^2A^2}{8} \times \frac{(q-\bar{q})^2}{q+\bar{q}} = \frac{4J^2b^2}{2a} \times \frac{J^2A^2}{8}$$

Thus

$$S = \frac{A^2b^2}{4a}$$

The three instruments therefore read

$$M = \frac{-J^2Ab}{a^2 - J^2b} \qquad S = \frac{A^2b^2}{4a} \qquad T = J^2Ab$$

as stated above.

(1) If the instruments are connected as shown in fig. (3) to act as a transient voltmeter it is clear that an instrument connected as shown at M will read

$$\int_{\infty}^{0} \iota dt = \int_{\infty}^{0} \left[\frac{EJe^{qt}}{r}\right]_1 dt$$

so that M in terms of A, a, and b will be determined by the same formula as before.

$$M = \frac{-J^2Ab}{a^2 - J^2b^2}$$

(2) If a condenser is connected from line to ground so that

$$Q = CE$$

$$i = C\frac{dE}{dt} = [CEJq2^{qt}]_1$$

If this current is passed through an instrument giving the integrated square value the reading S of the instrument will be given by the same formula as before

$$S = \frac{A^2b^2}{4a}$$

(3) If the current through such a condenser is passed through a current transformer and an instrument used to measure the secondary voltage

$$\frac{di}{dt} = C\frac{d^2E}{dt^2} = [CEJq^2\epsilon^{qt}]_1$$

then the integrated value will be $[EJq]_1 = T$ (ignoring calibration constant) and hence the reading T of the instrument will be given by the same formula as before

$$T = AJ^2b.$$

If we have two such waves for instance of current and E. M. F. each may be measured in the way proposed, giving values A_1 and q_1 for the first and A_0 and q_0 for the second. This will require six ballistic galvanometers for instance. If we connect a seventh as watt-second-meter as may easily be done, however, its readings can only be determined from $A_0A_1q_0$, q_1, by introducing a further parameter which we may call the generalized power-factor.

Suppose that the wave whose constants are A_0q_0 starts at $t = t_1$ instead of $t = 0$. Its expression will be

$$[A_0J\epsilon^{q_0t}\epsilon^{J\beta}]_1$$

and the watt-second-meter will read

$$W = \int_{-\infty}^{0} [A_0 J\epsilon^{q_0 t}\epsilon^{J\beta}]_1 [A_1 J\epsilon^{q_1 t}]_1 \, dt$$

$$= \frac{A_0 A_1}{4} \int_{-\infty}^{0} (J\epsilon^{q_0 t}\epsilon^{J\beta} + \bar{J}\epsilon^{q_0 t}\epsilon^{-J\beta})(J\epsilon^{q_1 t} + \bar{J}\epsilon^{\bar{q}_1 t}) \, dt$$

$$= \frac{A_0 A_1}{4} J^2 \int_{-\infty}^{0} [(\epsilon^{(q_0+q_1)t} - \epsilon^{(q_1+\bar{q}_1)t})\,\epsilon^{J\beta} - (\epsilon^{(\bar{q}_0+q_1)t} - \epsilon^{\overline{(q_0+q_1)}t})\,\epsilon^{-J\beta}] \, dt$$

Integrating

$$W = \frac{A_0 A_1 J^2}{4}\left[\left(\frac{1}{q_0+q_2} - \frac{1}{q_0+\bar{q}_1}\right)\epsilon^{J\beta} - \left(\frac{1}{\bar{q}_0+q_1} - \frac{1}{\overline{q_0+q_1}}\right)\epsilon^{-J\beta}\right]$$

or

$$\frac{W}{J^2} = C\epsilon^{J\beta} + \bar{C}\epsilon^{-J\beta}$$

for

$$C = \frac{\bar{q}_1 - q_1}{(q_0+q_1)(q_0+\bar{q}_1)}\,\frac{A_0 A_1}{4} \quad \bar{C} = \frac{q_1 - \bar{q}_1}{(\overline{q_0+q_1})(\bar{q}_0+q_1)}\,\frac{A_0 A_1}{4}$$

Put

$$x = \epsilon^{J\beta}$$

$$\frac{W}{J^2} = Cx + \frac{\bar{C}}{x}$$

$$0 = Cx^2 - \frac{W}{J^2}x + \bar{C}$$

$$x = \frac{\left(\frac{W}{J^2} \pm \sqrt{\frac{W^2}{J4} - 4C\bar{C}}\right)\bar{C}}{2C\bar{C}}$$

Now

$$C\bar{C} = \frac{-J^2 b_1^2}{[(a_0+a_1)^2 - J^2(b_0+b_1)^2]\,[(a_0+a_1)^2 - J^2(b_0-b_1)^2]}\,\frac{A_0^2 A_1^2}{16}$$

This can readily be calculated especially if $(b_0 - b_1)$ is small as is usually the case, and thus x determined and from it β. In fact since by definition x has the magnitude unity we have angle of x = angle of $\bar{C}$.

Giving

$$\tanh \beta = \frac{(a_0 + a_1)^2 + J^2(b_0^2 - b_1^2)}{2J^2(a_0 + a_1)\, b_0}$$

or approximately

$$\frac{a_0 + a_1}{2J^2 b_0} \qquad \text{if we neglect} \qquad (b_0^2 - b_1^2)$$

Now suppose the wave contains a second transient so that its expression is

$$[A_0 \epsilon^{J\beta}(\epsilon^{q_1 t} - \epsilon^{q_2 t})]_1$$

Suppose q_1, A_0 and β are determined as above

Let

$$M = \left[A_0\left(\frac{1}{q_1} - \frac{1}{q_2}\right)\epsilon^{J\beta}\right]_1$$

$$T = [A_0(q_1 - q_2)\, \epsilon^{J\beta}]_1$$

$$M - \left[\frac{A_0}{q_1}\epsilon^{J\beta}\right]_1 = M_0 = \left[\frac{A_0 \epsilon^{J\beta}}{q_2}\right]_1$$

$$T - [A_0 \epsilon^{J\beta} q_1]_1 = T_0 = [A_0 \epsilon^{J\beta} q_2]_1$$

Let

$$q_2 = R\epsilon^{J\alpha}$$

$$M_0 = \frac{A_0}{R}\cos(\text{h})(\beta - \alpha)$$

$$T_0 = A_0 R \cos(\text{h})(\beta + \alpha)$$

$$\frac{M_0 T_0}{A_0^2} = \cos(\text{h})(\beta - \alpha)\cos(\text{h})(\beta + \alpha)$$

From the formula

$$\cos(\text{h})\, A \cos(\text{h})\, \beta = \tfrac{1}{2}[\cos(\text{h})(A + \beta) + \cos(\text{h})(A - \beta)]$$

$$\frac{2M_0 T_0}{A_0^2} = (\cos(\text{h})\, 2\beta + \cos(\text{h})\, 2\alpha)$$

$$\frac{2M_0 T_0}{A_0^2} - \cos(\text{h})\, 2\beta = \cos(\text{h})\, 2\alpha \qquad \text{giving } \alpha$$

while

$$R = \frac{T_0}{A_0 \cos(\mathrm{h})(\beta + \alpha)} \qquad \text{gives } R.$$

Thus q_2 is determined. The notation cos(h) is used to show that the same argument applies to both types of transient.

Appendix II

Distorted transient waves

When one considers carefully what it is that gives us confidence in the vector method of treating alternating currents one cannot ignore the importance of our knowledge that distorted waves can be treated by the method of Fourier's series. If we did not know how to deal with distorted waves and feared that the slightest departure from the harmonic form might invalidate the vector method, our confidence in it would be materially diminished. Thus, though we may seldom make any actual use of Fourier's series, our knowledge that it could be used is an important constituent of our confidence so that it is important that we should have a similar knowledge in the case of transients. Obviously the use of series to represent transients is a very large subject even if we restrict ourselves to series made up of exponentials (Fourier's and Dirichlet's series). All that can be done here is to give one simple way of using a single series, there being every advantage in using Fourier's series owing to the many simple graphical and other methods which have been devised for calculating the coefficients. One way of utilising Fourier's series in connection with steady state phenomena may be illustrated as follows:

Let $E = 100 \sin \mathrm{wt} + 30 \sin 3\,\mathrm{wt}$ say

$$\mathrm{I} = 50 \sin\left(\mathrm{wt} + \frac{\pi}{4}\right) \text{ be the current and voltage in a given circuit}$$

The R.M.S. voltmeter reading will be 73.8, the ammeter reading 35.4 and the wattmeter reading 1770.

$$\text{Power} - \text{factor} = \frac{\text{Watts}}{\text{Voltamperes}} = .677 = \cos 47.4^\circ$$

Thus the equivalent sine waves of voltage and current will be:

$$E_1 = 104 \sin \mathrm{wt} \text{ and } I_1 = 50 \sin (\mathrm{wt} + 47.4)$$

$$= 50 (\sin \mathrm{wt} \cos 47.4 + \cos \mathrm{wt} \sin 47.4)$$

Subtracting these values from the true voltage and current we get

$$E - E_1 = E_r = -4 \sin wt + 30 \sin 3\,wt$$

$$I - I_1 = I_r = -50 \text{ x. } 677 \sin wt - 50 \text{ x. } 736 \cos wt + 50 \text{ x. } 707 \sin wt + 50 \text{ x. } 707 \cos wt.$$

$$= 50\,(.03 \sin wt - .03 \cos wt)$$

E_r and I_r we shall call the Residual Waves—they will in general have components of the fundamental frequency.

$$\text{Watts} = \int_0^{\frac{2\pi}{w}} EI\,dt = \int_0^{\frac{2\pi}{w}} (E_1 + E_r)(I_1 - I_r)\,dt$$

$$= \int_0^{\frac{2\pi}{w}} E_1 I_1\,dt + \int_0^{\frac{2\pi}{w}} (E_1 I_r + I_1 E_r + E_r I_r)\,dt$$

Since by definition the first of the above integrals yields the true watts the second must be zero.
Now

$$E_1 I_r + I_1 E_r + E_r I_r = EI_r + IE_r - E_r I_r = 0$$

From the above equation

$$EI_r + IE_r = E_r I_r$$

or in words. *The sum of the watts produced by the two original waves each multiplied by the appropriate residual is equal to the watts produced by the two residuals.* Clearly all the harmonics are to be found in the residuals.

If we suppose that we have an oscillographic record of a transient, and also the "equivalent wave" as determined by the instruments described above, giving us the constants A, a and b we can apply a similar method to the study of transients.

Plotting out our equivalent waves E_1 and I_1 on the same sheet and to the same scale as the oscillographic record we can determine the residual curves E_r and I_r by subtraction of ordinates. These will contain all the elements by which the oscillographic record differs from the second order wave.
We can now represent the wave as

$$E = E_1 \epsilon^{-at} \sin(h)\,bt + E_r$$

$$I = I_1 \epsilon^{-at} \sin(h)\,(bt - \beta) + I_2 \epsilon^{-a_2 t} \sin(h)\,(b_2 t - y) + I_r$$

E_r and I_r being represented by definite curves.

We may write

$$E_r = \epsilon^{-\alpha t}\epsilon^{\alpha t} E_r = \epsilon^{-\alpha t}E_s; \qquad I_r = \epsilon^{-\alpha t}\epsilon^{\alpha t}I_r = \epsilon^{-\alpha t}I_s$$

and plot the curves E_s and I_s. If the curve shows a continually increasing value the method will be difficult or impossible of application. If, however, it shows a decreasing value as it still will do in many of the most important practical cases we can expand E_s and I_s in Fourier series by any of the known methods.
Let the expansion be

$$E_s = \sum_n A_n^1 \cos b_n t + B_n^1 \sin b_n t$$

$$I_s = \sum_n A_n^{11} \cos b_n t + B_n^{11} \sin b_n t$$

Then the original current and voltage may be represented by

$$E = \epsilon^{-\alpha t}[E_1 (\sin (h)\, bt + \sum_n (A_n^1 \cos b_n t + B_n^1 \sin b_n t)]$$

$$I = I_2\epsilon^{-\alpha_2 t} \sin (h)\, (b_2 t + y) + \epsilon^{-\alpha t}[I_1 \sin (h)\, (bt + \beta) + \sum_n (A_n^{11} \cos b_n t - B_n^{11} \sin b_n t)]$$

The Fourier expansion can be used equally well for the residual curve of the hyperbolic transient as for that of a harmonic transient. We use the symbol $\sin(h)$ to indicate that either may be used.

Thus we have obtained an expression for many forms of distorted waves in terms involving only a single rate of decrement (apart from the "second transient" I_2 in the current wave). The arguments used above show that, as in the case of ordinary alternating currents

$$EI_r + IE_r = E_rI_r$$

Each term in the Fourier expansion when multiplied by $\epsilon^{-\alpha t}$ is a second order transient of the type studied above and the above methods therefore apply to each term individually. The effect of all the terms is the sum of the effects of each term separately. Hence we have a complete method for studying the effects of a distorted wave.

MATRICES IN ELECTRIC CIRCUIT THEORY

By Richard Stevens Burington[1]

1. **Introduction.** The recent appearance of a number of papers[2] on electric circuit theory in which matric theory has been used makes it appropriate to discuss certain of these applications and their significances in the light of the theory of matrices.

In some of these papers the language of matric theory has been used in varying degrees of elegance and generality while in others the notation of the particular application has been used and the matric theory generality has not been recognized.

In the present paper, a short synopsis of certain manifestations of matric algebra which have arisen in electric circuit theory are discussed, though no attempt is made to give more than an outline of certain phases of the subject. The first part is devoted to purely mathematical considerations, starting with a discussion of certain types of equality relationships in matric algebra and their accompanying theories. The second part is concerned with a discussion of these principles in electric circuit theory.

The introduction of the concept of direct sum into the electrical theory given here should be of particular interest in clearing up certain difficulties in the literature.

PART I. THE THEORY OF MATRICES[3]

2. **Matric Rings.** A *total matric ring* (or *algebra*) of order n over a ring R is a mathematical system M whose elements are the arrays

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = (a_{rs}) \tag{2.1}$$

[1] Mathematics Department, Case School of Applied Science, Cleveland, Ohio.

[2] See references (6)–(9), (12)–(19), (28), (33)–(35), (42) at end of paper.

[3] See the Symposium lecture of C. C. MacDuffee, "Matrices with Elements in a Principal Ideal Ring," Bulletin of the American Mathematical Society,

where the a_{rs} belongs to R, subject to the following operations and postulates: (1) Two arrays $A = (a_{rs})$ and $B = (b_{rs})$ are equal if $a_{rs} = b_{rs}$ for every r and s. (2) The *sum* is defined by

$$A + B = (a_{rs} + b_{rs}). \tag{2.2}$$

The set whose elements are all zero will be denoted by O. (3) The *product* is defined by

$$AB = \left(\sum_{j=1}^{n} a_{rj} b_{js} \right). \tag{2.3}$$

The product is unique; $OA = AO$ for every A; multiplication in M is associative; and distributive with respect to addition, though not commutative in general.

Each of the arrays of M is called a *matrix*.

If R has a principal unit 1, then $I = (\delta_{rs})$ is called the *identity matrix* of M, where δ_{rs} is 1 if $r = s$ and 0 if $r \neq s$.

As in the case of a field, it is desirable to have associated with every A of an algebra M a number $\Theta(A)$ which will serve the purpose of an absolute value. This may be done by associating a scalar function $\Theta(A)$ of the elements a_{rs} of A such that $\Theta(A)$, for every A, is a non-constant rational integral function of the a_{rs} of lowest degree such that

$$\Theta(AB) = \Theta(A) \cdot \Theta(B). \tag{2.4}$$

It is known that such a function $\Theta(A)$ exists and is equal to

$$\Theta(A) \equiv \sum_{h_1, \cdots, h_n = 1}^{n} \epsilon_{h_1 \cdots h_n} \cdot a_{1h_1} \cdot a_{2h_2} \cdots a_{nh_n} \equiv d(A), \tag{2.5}$$

where the sum is over all permutations $(h_1, \cdots, h_n)$ of $(1, 2, \cdots, n)$, and $\epsilon_{h_1 \cdots h_n}$ is 1 or -1 according as the permutation is even or odd. $\Theta(A)$ is called the *determinant* of A.

The elementary properties of determinants are well known.[4]

If $S^T = S$, S is called *symmetric*; if $Q^T = -Q$, Q is called *skew*.

Theorem 2.1. *Every matrix A over a ring, in which $2x = a$ is solvable for every a, is uniquely expressible as a sum of a symmetric and a skew matrix.*

Aug. 1933, pp. 564–584, which lends itself in a most appropriate manner to the present development. Also, see C. C. MacDuffee, "Theory of Matrices," Berlin, 1933.

[4] See references (4), (32), (39).

3. **Types of Equality in Matric Theory and their Significances in Circuit Theory.** The notion of equality lies at the very basis of nearly every mathematical system and hence is of great importance in any mathematical analysis or formulation of a physical theory. The abstract formulation of this notion of equality has been discussed by a number of writers[5] and may be embodied in the following four postulates. The relation denoted by $A = B$ is an equality relation if it is:

(1) *Determinative* (either $A = B$ or $A \neq B$),

(2) *Reflexive* ($A = A$),

(3) *Symmetric* ($A = B$ implies $B = A$),

(4) *Transitive* (If $A = B$ and $B = C$, then $A = C$).

Each such definition of equality constitutes a separation of a set of elements into classes. In the considerations to follow, this separation will be seen to be of considerable significance, when the elements involved are given physical meanings.

It has become to be recognized that the richness of matric theory is due in a large part to the number of non-isomorphic types of equality which can be defined, each type of equality having associated with it an extensive theory. In this section a few examples of types of equality are given which arise in various branches of electrical theory and which when properly abstracted lead to the following types of equality in matric theory.

(1) A is a *left associate* of B if there exists a unimodular matrix U (U is *unimodular*, or a *unit matrix*, if there exists a matrix U' such that $UU' = I$) such that $A = UB$, written $A \overset{L}{=} B$, where the matrices belong to a ring[6] with elements in a principal ideal ring P.

(2) Likewise, A is a *right associate* of B, $A \overset{R}{=} B$, if there exists a U such that $A = BU$.

(3) A is *equivalent* to B, $A \overset{E}{=} B$, if there exists two unimodular matrices U and V such that $A = UBV$.

(4) A is *congruent* to B, $A \overset{c}{=} B$, if there exists a unimodular matrix U such that $A = U^{T}BU$, where U^{T} is the transpose of U, that is, U^{T} is the matrix obtained from U by changing rows to columns.

[5] See references (38), (39), (42).

(5) A is *m-affine congruent*[6] to B, $A \overset{me}{=} B$, if there exists a unimodular matrix U of the type

$$(3.1) \qquad U = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & & \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ u_{m+1,1} & \cdots & u_{m+1,m} & & u_{m+1,m+1} & \cdots & u_{m+1,n} \\ \cdot & & & & & & \\ \cdot & & & & & & \\ u_{n,1} & \cdots & u_{n,m} & & u_{n,m+1} & \cdots & u_{n,n} \end{bmatrix}$$

such that $A = U^{T}BU$.

(6) A is similar to B, $A \overset{s}{=} B$, if there exists a unimodular matrix U such that $A = U^{-1}BU$, where U^{-1} is the inverse of U, that is, a unit matrix such that $U^{-1}U = I$.

Similar definitions hold for m-affine equivalence and similarity.

Other types of equality have been defined, but since no use will be made of them here, they will not be listed.

Relations (1) to (6) obey the four postulates listed above and hence are (abstractly) relations of equality.

Evidently, congruence and similarity are instances of equivalence.

In later sections of this paper these types of equality will be referred to and will appear in distinctly different physical settings.

4. **Theory of Elementary Matrices.** An *elementary operation* or *transformation* upon the rows of a matrix A is a transformation upon the rows of A of one or more of the following types:

(1) The interchange of two rows.

(2) The addition to the elements of a row of K times the corresponding elements of another row, K being in P.

(3) The multiplication of the elements of a row by a unit of P.

Each of these elementary operations can be affected by multiplying A on the left by a certain *elementary matrix*, a unimodular matrix obtained by performing the elementary operation under consideration upon the identity matrix I.

[6] See reference (9).

If B is derived from A by a finite number of elementary transformations upon the rows of A, $B \overset{L}{=} A$. Similar elementary operations upon the columns of a matrix may be accomplished by multiplying the matrix on the right by an elementary matrix, in which case $B \overset{R}{=} A$.

A generalization of a theorem of **Hermite's**, of considerable importance and utility, has been given by MacDuffee:[7]

Theorem 4.1. *Every matrix A with elements in P is the left associate of a matrix having 0's above the main diagonal, each diagonal element lying in a prescribed system of non-associates, and each element below the main diagonal lying in a prescribed residue system modulo the diagonal element above it. If, furthermore, any diagonal element is 0, all the elements of its rows can be made 0. (Hermite's Normal Form.)*

5. **Linear Dependence Rank.** The notion of linear dependence is of considerable importance, and is essential to the understanding of many theorems in mathematics as well as physics.

The quantities $x_1, \cdots, x_n$ (numbers, n sets of m constants, polynomials, vectors, etc.) are said to be *linearly dependent* in the field F if there exists numbers $c_1, \cdots, c_n$ in F, not all zero, such that

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = 0 .$$

If no such numbers c_i exist, $x_1, \cdots, x_n$ are said to be *linearly independent* with respect to F. Thus x_1 is linearly dependent if $x_1 = 0$ and linearly independent if $x_1 \neq 0$.

If A is a rectangular array of m rows and n columns with elements in a field F, the rank ρ of A is defined to be the order of a minor square array of A of maximum order whose determinant is not zero. If $m = n$, A is square and $n - \rho$ is called the *nullity*[8] of A.

Theorem 5.1. *If A, B, C are matrices with elements in a field F and of order n*

$$\rho(AB) + \rho(BC) \leqq \rho(B) + \rho(ABC) ,$$

$$\rho(AB) \leqq \rho(A) \quad \text{and} \quad \rho(AB) \leqq \rho(B) .$$

A theorem[9] of considerable generality and utility is

Theorem 5.2. *Let A be a square array of order n, with elements in a*

[7] See reference (38).
[8] See reference (37).
[9] See reference (39), pp. 10–11.

field F, of nullity κ and rank ρ. There exist exactly κ linearly independent linear relations among the rows (columns) of A, and conversely.

In part II of this paper, many theorems involving the use of the notion of nullity and rank are encountered.

6. **Linear Systems.**[10] The set of linear equations (or transformations)

$$y_r = \sum_{s=1}^{n} a_{rs} x_s \qquad (r = 1, \cdots, n) \tag{6.1}$$

may be written

$$\{y\} = A\{x\} \tag{6.2}$$

where

$$\{y\} = \begin{Bmatrix} y_1 \\ \vdots \\ y_n \end{Bmatrix}, \qquad \{x\} = \begin{Bmatrix} x_1 \\ \vdots \\ x_n \end{Bmatrix}$$

are column arrays or vectors, and A is a matrix of order n, the "row by column" rule of multiplication being used. $\{y\}$ is said to be a linear vector function of $\{x\}$.

If A is non-singular (*i.e.* of rank n), then (6.2) has a unique solution

$$\{x\} = A^{-1}\{y\} \tag{6.3}$$

where A^{-1} is the *inverse* of A.

Equation (6.3) is sometimes written

$$x_r = \sum_{s=1}^{n} a^{rs} y_s \qquad (r = 1, \cdots, n) \tag{6.4}$$

where

$$(a^{rs}) \equiv A^{-1} \equiv (A_{rs}/d(A)), \qquad d(A) \neq 0, \tag{6.5}$$

A_{rs} is the *co-factor* of a_{rs} in A, and $(A_{rs}) = A^A$ is the *adjoint* of A, obtained from A by replacing every a_{rs} by A_{sr}.

Various interpretations have been placed upon (6.2) and (6.3).

Theorem 6.1. *A linear transformation (or vector function) over a field F with a matrix B, $\{x\} = B\{z\}$, replaces a system of linear forms (or vectors), $\{y\} = A\{x\}$, with matrix A, in F, by a system of linear forms (or vectors), $\{y\} = AB\{z\}$, with matrix AB. If C be the matrix of a third*

[10] See references (4), (39), (53).

linear transformation (or vector function), $\{z\} = C\{w\}$, then $A(BC) = (AB)C$ is the matrix of the resulting system, $\{y\} = ABC\{w\}$.

Theorem 5.2 can be translated into theorems on the linear dependence of linear forms, etc. For example,[11]

Theorem 6.2. *Given m homogeneous linear equations in n unknowns with coefficients in F,*

$$\sum_{s=1}^{n} a_{rs} x_s = 0\,, \qquad (r = 1, \cdots, m)\,. \tag{6.6}$$

Let ρ be the rank of the array $A = (a_{rs})$. Then (6.6) has $(n - \rho)$ linearly independent solutions in F, while every other solution is linearly dependent on them, (i.e., ρ equations may be selected from (6.6) so that this matrix has a non-vanishing ρ-rowed determinant, and these ρ equations determine uniquely ρ of the unknowns as homogeneous linear functions with coefficients in F, of the remaining $(n - \rho)$ unknowns; for all values of the later, the expressions for the ρ unknowns satisfy (6.6)).

Theorem 6.3. *Given the non-homogeneous linear equations with coefficients in F*

$$y_r = \sum_{s=1}^{n} a_{rs} x_s \qquad (r = 1, \cdots, m) \tag{6.7}$$

of array $A = (a_{rs})$. Let r be the rank of A and ρ the rank of B, the augmented array derived from A by annexing the column $\{y\}$. If $\rho > r$, (6.7) are inconsistent. If $\rho = r$, certain r of the equations determine uniquely r of the unknowns as linear functions of the remaining $(n - r)$ unknowns; for all values of the latter, the expressions for these r unknowns satisfy also the remaining $(m - r)$ equations.

7. **Equivalence.**[12] Two matrices A and B with elements in a principal ideal ring P are said to be *equivalent* if two unimodular matrices U and V exist with elements in P such that $A = UBV$.

Theorem 7.1. *Every matrix A of rank ρ with elements in a principal ideal ring P is equivalent to a diagonal matrix $[h_1, h_2, \cdots, h_\rho, 0, \cdots, 0]$ where $h_i \neq 0$ and $h_i \mid h_{i+1}$. (Smith's normal form.)*

Since every polynomial domain (λ-matrices) with one indeterminate of a field is a principal ideal ring, the units being the non-zero numbers of the field, this theorem will be of interest in part II.

[11] See reference (20), p. 62.

[12] See references (38), (39), (51), (53).

Suppose

$$h_i = p_1^{e_{i1}} p_2^{e_{i2}} \cdots p_k^{e_{ik}}$$

is factored into primes. Since $h_i \mid h_{i+1}$, a sequence

$$e_{ni} \geqq e_{n-1,i} \geqq \cdots \geqq e_{1i} \qquad (i = 1, \cdots, k)$$

is formed by the exponents of each prime factor. Such of the powers $p^{e_{ij}}$, not units in P, are called *elementary divisors* of A. They are defined up to a unit factor. (See §21.)

The quotients $h_1 = d_1$, $h_2 = d_2/d_1$, $h_3 = d_3/d_2$, $\cdots$ are called the *invariant factors* of A, and are invariant under equivalence, where $d_i = h_1 h_2 \cdots h_i$.

The exponents e_{ij} of the elementary divisors are *invariants of structure* and

$$\begin{matrix} e_{11} & e_{21} & \cdots & e_{j1} \\ e_{12} & e_{22} & \cdots & e_{j2} \\ \cdot & \cdot & \cdots & \cdot \end{matrix}$$

is known as the *Segre characteristic* of a matrix A.

Theorem 7.2. *The g.c.d. d_i of the i-rowed minor determinants of A divides the g.c.d. d_{i+1} of the $(i+1)$-rowed minor determinants of A.*

Theorem 7.3. *A is equivalent in P to B if and only if A and B have the same elementary divisors (or invariant factors). (If A and B have elements in a commutative field F, $A \overset{E}{=} B$ if and only if they have the same rank and $A \overset{E}{=} [1, \cdots, 1, 0, \cdots, 0]$).*

Theorem 7.4. *The pairs of matrices A_1, A_2 and B_1, B_2, A_1 and B_1 non-singular, are equivalent if and only if $A_1\lambda + A_2$ and $B_1\lambda + B_2$, in the polynomial domain $P(\lambda)$, have the same invariant factors (or elementary divisors).*

Theorem 7.5. *Two non-singular pairs of matrices A_1, A_2 and B_1, B_2, in a field F, are equivalent if and only if they have the same invariant factors.*

8. **Congruence.**[13] A is said to be *congruent* to B if there exists a unimodular matrix U such that $A = U^T B U$, where all matrices have elements in a principal ideal ring P.

By theorem 2.1, for every matrix A, $2A = S + Q$.

If $2A_1 \overset{c}{=} 2A_2$, then $S_1 \overset{c}{=} S_2$ and $Q_1 \overset{c}{=} Q_2$. Thus the properties of

[13] See references (38), (39).

symmetry and skewness are invariant under congruence. In fact, the invariants of S and Q are invariants of $2A = S + Q$.

The problem[13] of the congruence of symmetric matrices is one of great difficulty, though the corresponding problem for skew matrices has been completely solved for matrices with elements in P.

Since in part II, symmetric matrices over a field occur most frequently no further mention will be made of the properties of skew matrices.

Theorem 8.1. *Every symmetric matrix of rank ρ with elements in a field F, not of characteristic 2, is congruent in F with a diagonal matrix $[g_1, \cdots, g_\rho, 0, \cdots, 0]$, $g_i \neq 0$. (Each g_i can be determined as a rational function of the elements of A_i alone, where A_i is the principal minor matrix of order i in the upper left hand corner of A, and $d(A_i) \neq 0$.)*

Theorem 8.2. *In an ordered field, the rank $\rho = p + n$ and the signature $\sigma = p - n$ are integer invariants of a symmetric matrix A under the congruence $P^\tau AP$, where p is the number of positive terms in the normal form $[g_1, \cdots, g_\rho, 0, \cdots, 0]$ and n is the number of negative terms.*

Theorem 8.3. *In an algebraically closed field, two symmetric matrices are congruent if and only if they have the same rank ρ.*

Theorem 8.4. *A necessary and sufficient condition that two real symmetric matrices be congruent in the real field R is that they have the same rank and signature.*

A symmetric matrix A of order n in an ordered field is called *positive definite* if $\rho = \sigma = n$, and negative definite if $\rho = -\sigma = n$. It is *semi-definite* if $\rho = \sigma < n$, or $\rho = -\sigma < n$.

Theorem 8.5. *In a commutative field, if $A = P^\tau BP$, then*

$$d(A) = d(B) \cdot [d(P)]^2$$

is a relative invariant of weight 2 of B under $A \overset{c}{=} B$.

Theorem 8.6. *If A and B are symmetric with elements in an algebraically closed field, and if $A \overset{E}{=} B$, then $A \overset{c}{=} B$ in the field.*

Theorem 8.7. *Two pairs of symmetric matrices (A, B) and (A_1, B_1) in an algebraically closed field are congruent if and only if they are equivalent.* (See Theorems 7.3, 7.4.)

Theorem 8.8. *If A and B are real symmetric matrices and A is positive definite, there exists a non-singular matrix P such that $I = P^\tau AP$ and $D = P^\tau BP$, where I is the identity matrix and D the diagonal matrix $[\lambda_1, \cdots, \lambda_n]$, λ_i being the roots (all real) of $d(A - \lambda B) = 0$.*

The following theorem is of interest.

Theorem 8.9. *Every matrix A satisfies its characteristic equation*

$$d(\lambda I - A) = 0.$$

If the quadratic form $F = \sum_{i,j=1}^{n} b_{ij} x_i x_j$ be subjected to the linear transformations $\{x\} = P\{x'\}$, $d(P) \neq 0$, then the matrix B of form F becomes $A = P^T B P$.

Two quadratic forms F_1 and F_2 in a field F are said to be *equivalent* in F, if and only if there exists a transformation P which transforms F_1 into F_2.

Theorem 8.10. *Two quadratic forms in F are equivalent if and only if their matrices are congruent.*

It is now evident how the theorems on congruence may be translated into theorems on quadratic forms. (See §17.)

A theorem of interest in connection with the solution of polynomial equations is the Borchardt-Jacobi Theorem:[14]

Theorem 8.11. *The number of distinct roots $\xi_1 \cdots \xi_n$ of*

$$E \equiv a_0 x^n + n a_1 x^{n-1} + \cdots + a_n = 0, \qquad a_0 \neq 0$$

with real coefficients is equal to the rank ρ of the discriminant matrix

$$M \equiv (S_{r+s-2}), \qquad S_\lambda \equiv \sum_{j=1}^{n} \xi_j^\lambda .$$

The number of distinct real roots is equal to the signature σ of M.

9. **m-affine congruence.**[15] Two matrices A and B are said to be *m-affine congruent* if there exists a unimodular matrix T of type

$$T = \left[\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & & \cdot \\ \cdot & & & 0 & \cdot & & \cdot \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \hline t_{m+1,1} & \cdots & & t_{m+1,m} & t_{m+1,m+1} & \cdots & t_{m+1,n} \\ \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \\ t_{n,1} & \cdots & & t_{n,m} & t_{n,m+1} & \cdots & t_{n,n} \end{array}\right] \tag{9.1}$$

such that

$$A = T^T B T, \tag{9.2}$$

where all the matrices have elements in a principal ideal ring P.

[14] See references (3), p. 168; (8).

[15] See references (6), (7), (9). Also §18.

In this section only symmetric matrices with elements in a field F are considered. If $m = 0$, the congruence is the usual one discussed in §8.

The corresponding theorems on quadratic forms may be translated from those given on matrices in the usual manner.

It has been shown that[15]

Theorem 9.1. *If* $H \equiv A_{r_1 \cdots r_t}^{s_1 \cdots s_t}$ *is matrix* A *(in ring* R*) with rows* $r_1 \cdots r_t$ *and columns* $s_1 \cdots s_t$ *deleted,* $(r_i, s_i, \leqq m)$*, then* H *is an invariant matrix of* A *under* T *(in* F*), that is,*

$$\bar{A}_{r_1 \cdots r_t}^{s_1 \cdots s_t} = T_{r_1 \cdots r_t}^{T} \cdot A_{r_1 \cdots r_t}^{s_1 \cdots s_t} \cdot T_{s_1 \cdots s_t}. \tag{9.3}$$

The determinants $d(H)$ *are relative invariants, their ratios absolute invariants (parameters), and the ranks* $\rho(H)$ *integer invariants.*

Every matrix A can be reduced to certain canonical forms $c_1, c_2, \cdots$ according to the ranks (and signatures if F is ordered) of the invariant matrices H of A.

Theorem 9.2. *A necessary and sufficient condition for the* m*-affine congruence of two matrices* A *and* B*, whose elements belong to the real field* R *is that their invariant matrices have the same ranks and signatures, respectively, and that their parameters, in case they appear in form* C_i *under consideration, be identical. If the field is algebraically closed, the preceding statement holds without signatures.*

Two quadratic forms in a field F are m-affine equivalent if and only if their matrices are m-affine congruent in F.

C. Goffman[16] has considered the problem of the m-affine congruence of pairs of matrices in a field.

The relation of this type of equality to circuit theory is discussed in Part II.

10. **Similarity.**[17] Two matrices A and B with elements in a principal ideal ring $\mathfrak{S}$ are called *similar* if there exists a unimodular matrix P such that $A = P^{-1}BP$. Similar matrices may be interpreted as representing the same transformation referred to different bases. The theorems of this section may be so interpreted. (See §19.)

Theorem 10.1. *The coefficients of the characteristic function* $d(\lambda I - A)$ *of* A *are similarity (absolute) invariants.*

Corollary 10.1. *If* $A \overset{S}{=} B$ *in an algebraically closed field, the characteristic roots of* B *coincide with those of* A*, and each has the same multiplicity in* B *as in* A*.*

[16] See reference (24).

[17] See references (38) (39).

Theorem 10.2. *A necessary and sufficient condition that* $A \overset{S}{=} B$, *in* F, *is that in the polynomial domain* $F(\lambda)$, $(I\lambda - A)$ *and* $(I\lambda - B)$ *have the same invariant factors.*

11. **Direct Sum.**[18] The *direct sum* of two matrices A and B of orders α and β, whose elements belong to a ring is defined to be

$$A + B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

and is of order $(\alpha + \beta)$. Among the many properties, the following are of particular interest here. (See §20.)

Theorem 11.1. *In* P, $\rho(A + B) = \rho(A) + \rho(B)$, $\rho(A)$ *being the rank of* A.

Theorem 11.2.
$$(A + B) + C = A + (B + C).$$
$$(A_1 + A_2) + (B_1 + B_2) = (A_1 + B_1) + (A_2 + B_2).$$
$$(A_1 + B_1)(A_2 + B_2) = A_1A_2 + B_1B_2.$$
$$(A + B)^T = A^T + B^T.$$
$$(A + B)^{-1} = A^{-1} + B^{-1}.$$
$$d(A + B) = d(A) \cdot d(B).$$

Theorem 11.3. *The characteristic roots of* $A + B$ *are the characteristic roots of* A *together with those of* B.

Theorem 11.4. *In a commutative ring, if* $QAQ_1 = A_1$ *and* $RBR_1 = B_1$, *then* $(Q + R)(A + B)(Q_1 + R_1) = A_1 + B_1$.

12. **Further Results in Matric Theory.** The literature[19] on matric theory is very extensive. The earlier sections of this paper have dealt only with certain elementary topics which appear to be fundamental in the matric development of circuit theory, no attempt having been made to cover the field. For a more extensive treatment, reference should be made to the various papers and works in the field. A short bibliography is listed at the end of this paper.

PART II. ELECTRICAL CIRCUIT THEORY

13. **Introduction.** In this section indications are given as to the manner in which certain phases of the theory of electrical circuits may be developed in the language of matric theory. The various types of equalities and compositions discussed in Part I are brought out with emphasis upon their physical significances.

14. **Linear Networks.**[20] A *linear network* is an electrical network

[18] See references (39), p. 81.

[19] See references at end of paper.

[20] See references at end of paper, particularly (25), (30), (47).

consisting of a system of resistances (R), inductances (L), and capacitances (C), interconnected with sources of electromotive forces (e.m.f.'s), and where the R, L, C's are assumed to be constants.

Consider the $2m$-terminal n-mesh linear electrical network containing (lumped) resistances, inductances and elastances. Let $e_1, \cdots, e_m$ be the (real) e.m.f.'s impressed on terminal pairs $1, \cdots, m$ respectively; q_s and i_s, be the instantaneous (real) charges and (real) currents, respectively, in mesh s; R_{st}, L_{st}, D_{st}, $s \neq t$, (real numbers) be the (lumped) circuit parameters (resistance, industance and elastance, respectively) mutual to meshes s and t; and R_{ss}, L_{ss}, D_{ss}, the total circuit (real) parameters of mesh s, i.e., the total resistance, inductance and elastance of mesh s, $(s, t = 1, \cdots, n)$. The meshes are so chosen that mesh s is the only one which passes through the terminal pairs s, $(s = 1, \cdots, m)$.

By Kirchoffs' laws, an equation for each circuit of the network may be written. The complete differential equations for the network of n meshes are

$$A\{i\} = \{e\} \tag{14.1}$$

where

$$a_{st} = R_{st} + L_{st}p + D_{st}p^{-1}, \qquad p = d/dt,$$

the p being the usual derivative operator and p^{-1} its inverse. $A \equiv (a_{rs})$ is a *network matrix* and $\{i\} = \{i_1, \cdots, i_n\}$, $\{e\} = \{e_1, \cdots, e_m, 0, \cdots, 0\}$ are vectors or one-column arrays. In this section, it is assumed that the mutual impedances a_{rs} are (bi-lateral) reciprocal; i.e., that A is a symmetric matrix, and hence $R = (R_{st})$, $L = (L_{st})$, $D = (D_{st})$ are symmetric.

The matrix A depends not only upon the circuit parameters but also upon the particular agreement made upon the selection of the paths forming the meshes, the number of these meshes, and the direction in which the currents in these meshes are assumed to take.

The various network matrices B derivable from a given network of matrix A are all derivable one from another by elementary transformations. Thus, Kronecker's theory of elementary matrices enters the theory (§4) and the equality relationships of associates enter. If (14.1) be subject to a change in agreement, etc., represented by the transformations $\{i\} = Q\{i'\}$, then $AQ\{i'\} = \{e\}$ and $A \overset{R}{=} B \equiv AQ$.

If both currents and e.m.f's are subject to non-singular transformations[21]

$$\{i\} = Q\{i'\} \qquad \text{and} \qquad \{e\} = P\{e'\}, \tag{14.2}$$

[21] See reference (35).

then (14.1) becomes

$$AQ\{i'\} = P\{e'\} \qquad \text{or} \qquad B\{i'\} = \{e'\}, \tag{14.3}$$

where

$$B = P^{-1}AQ. \tag{14.4}$$

Thus, $A \overset{E}{=} B$. If $P = Q$, then $A \overset{S}{=} B$. (See §7, §10).

15. **Steady-State Solution.**[22] The general solution of (14.1) is the sum of the particular or *steady-state* solution and the complementary or *transient* solution. (A method well known to electrical engineers is used here for purposes of exposition.)

Let the e.m.f's be given by

$$e_\mu = E_\mu e^{j\omega t}, \quad (\mu = 1, \cdots, m), \quad e_\mu = 0, \quad (\mu = m+1, \cdots, n), \quad j^2 = -1$$

or

$$\{e\} = \{E\}e^{j\omega t}, \tag{15.1}$$

where E_μ is the complex-root-mean square voltage amplitude for pair μ. The real part of e_μ is the actual e.m.f., ϵ_μ.

Suppose

$$\{i\} = \{I\}e^{j\omega t} \tag{15.2}$$

is a solution of (14.1). Then

$$B\{I\} = \{E\}, \tag{15.3}$$

where

$$B = (b_{rs}), \qquad b_{rs} = b_{sr} = R_{rs} + L_{rs}\lambda + D_{rs}\lambda^{-1}, \qquad \lambda = j\omega.$$

If $d(B) \neq 0$,

$$\{I\} = B^{-1}\{E\}. \tag{15.4}$$

If $B^A = (B_{\mu k})$ is the adjoint of B, then $Z_{\mu k} = d(B)/B_{\mu k}$ is the *generalized network impedance*, a *transfer impedance* if $\mu \neq k$, a *driving-point impedance* if $\mu = k$. (See §6.) The steady-state solution of (14.1) is (15.2) with $\{I\}$ given by (15.4). B^A is symmetric if B is symmetric. $I_k/I_s = B_{\mu k}/B_{\mu s} = b_{\mu s}/b_{\mu k}$.

The *Superposition Principle* and the *Reciprocal Theorem* follow directly from (15.4). The proof of the Generalized Reciprocal Theorem is very concise.

[22] See references (10), (25), (30); (41), pp. 246–313; (47).

Theorem 15.1. *If a set* $\{E'\}$ *of e.m.f.'s, all of the same frequency acting in the m branches of an invariable network, produce a current distribution* $\{I'\}$, *and a second set* $\{E''\}$ *of the same frequency produce a second current distribution* $\{I''\}$, *then* $(E')\{I''\} = (E'')\{I'\}$, *i.e.,*

$$\sum_{i=1}^{m} E'_i I''_i = \sum_{i=1}^{m} E''_i I'_i.$$

Proof. By (15.4) $\{I'\} = B^{-1}\{E'\}$ and $\{I''\} = B^{-1}\{E''\}$ or $(I'') = (E'')(B^{-1})^T$. $[(I)$ is a row vector $(I_1, \cdots, I_m)]$. Multiply by the transpose (E'') of $\{E''\}$. Then $(E'')\{I''\} = (E'')B^{-1}\{E'\} = (I'')\{E'\}$, for $(I'') = (E'')B^{-1}$, $(B^{-1})^T = B^{-1}$, since B is symmetric.

If several different e.m.f.'s of different frequencies ω_s are simultaneously impressed, each of the type $e_{r,s} = E_{r,s}e^{j\omega_s t}$, $e_{r,s}$ being the e.m.f. impressed in mesh r of frequency ω_s, then

$$\{i\} = (I_{r,s})\{e^{j\omega_s t}\} = \left\{\sum_{s=1}^{\mu} I_{r,s}e^{j\omega_s t}\right\} \tag{15.5}$$

is the steady-state solution, where $(I_{r,s})$ is a rectangular array and $I_{r,s}$ is the current amplitude in response to frequency ω_s in mesh r, $i_{r,s} = I_{r,s}e^{j\omega_s t}$.

In connection with (14.1) and its solution, it may happen that (15.3) are insufficient to yield a unique solution due to the fact that all the independent relationships that can be expressed by Kirchoffs' laws have not been utilized, and/or some of the equations may be linearly dependent upon certain others; i.e., a linear dependence relation may exist between the e's, etc. Such situations arise in the theory of ideal transformers. Here, the theories of §5, §6 are useful; e.g., the rank r of B is the number of equations determining uniquely r of the currents $\{I\}$ as linear functions of the remaining $(n - r)$ I's and $(n - r)$ is the number of linear independent mesh currents.

16. **Transient Solution.**[23] The transient solution of (14.1) is the solution of

$$A\{i\} = \{0\}. \tag{16.1}$$

Let a solution of (16.1) be

$$i_\mu = J_\mu^{(\nu)} e^{p_\nu t} \qquad (\mu = 1, \cdots, n)$$

[23] See references (10), (11), (25); (41), pp. 246–313.

or

$$\{i\} = \{J^{(\nu)}\}\, e^{p_\nu t}. \tag{16.2}$$

Then

$$C^{(\nu)}\{J^{(\nu)}\} = \{0\} \tag{16.3}$$

where

$$C^{(\nu)} = (C^{(\nu)}_{jk}), \qquad C^{(\nu)}_{jk} = R_{jk} + L_{jk}\, p_\nu + D_{jk}\, p_\nu^{-1}.$$

If (16.2) is valid, then so must (16.3). For each *mode* p_ν, there is a set of equations (16.3). $\{0\} = \{J^{(\nu)}\}$ is a trivial solution of (16.3). In (16.2), the p_ν are undetermined. By §5, for (16.3) to have a non-trivial solution, the *determinantal equation*, of degree $\tau \leqq 2n$,

$$d(C^{(\nu)}) \equiv D(p_\nu) = 0, \tag{16.4}$$

must have solutions p_ν. The p_ν are called the *natural modes.*

The nullity $\kappa = (n - r)$ of $C^{(\nu)}$ is the number of independent solutions of (16.3), every other solution being linearly dependent upon them, r being the rank of $C^{(\nu)}$. (See §6.)

A solution X of $C^{(\nu)}X = 0$ of nullity $(n - r)$ is called a *complete solution.*[24] Then, if $\{x\}$ be arbitrary,

$$\{J^{(\nu)}\} = \beta^{(\nu)}\{x\} \tag{16.5}$$

is the most general solution of (16.3), $\beta^{(\nu)}$ being a complete solution of $C^{(\nu)}\beta = 0$.

The solution of (16.1) for p_ν is

$$\{i\} = \{J^{(\nu)}\}\, e^{p_\nu t} = \beta^{(\nu)}\{x\}\, e^{p_\nu t}. \tag{16.6}$$

If all the natural modes p_ν are distinct, then the transients are

$$\{i\} = J\{e^{p_\nu t}\}, \tag{16.7}$$

$J \equiv (J^{(\nu)}_k)$ being a rectangular array of k rows and ν columns.

Many theorems can be stated concerning the modes of the network. The number n of degrees of freedom of a network is the number of independent meshes, that is, the least number of meshes by which a network may be specified.[25]

Theorem 16.1. *If the network A is passive (that is, R, L, D are positive definite), the real modes p_ν are negative, and the complex roots always occur*

[24] See reference (39), pp. 10–11.

[25] See references (25); (43), p. 7.

in conjugate pairs with their real portions negative. The maximum number of modes that a network may contain is twice the number of independent meshes. $\tau \leqq 2n$, *is the number of independent integration constants of* (16.1), *and the number of independent modes of* (16.1)

Theorem 16.2.[26] *Let* M *be the discriminant matrix of* (16.4). *The signature* σ *of* M *is the number of distinct real natural modes, and the rank* ρ *of* M *is the number of distinct natural modes.* ρ *and* σ *are integer invariants of* M *under transformations* (14.5). (See Theorem 8.11.)

While (16.7) contains $2n^2$ amplitudes $J_k^{(r)}$, only the amplitudes for one mesh current need be considered as the integration constants of system (16.1). In electrical networks, r is either n or $(n-1)$. From (16.3) for mode v, and any $i = 1, \cdots, n$

$$J_k^{(r)} = \Gamma_{ik}^{(r)} G^{(v)} \qquad (k = 1, \cdots, n, v = 1, \cdots, 2n) \tag{16.8}$$

where $\Gamma^{(v)} = (\Gamma_{ik}^{(r)})$ is the *adjoint* matrix of $C^{(v)}$, and $G^{(v)}$ is an arbitrary constant.

$$J_k^{(r)} = \Gamma_{ik}^{(r)} J_s^{(r)} / \Gamma_{is}^{(r)} \qquad (k = 1, \cdots, n, v = 1, \cdots, 2n). \tag{16.9}$$

Thus, all the J's are expressible in terms of those for the s^{th} mesh, and may be considered the true integration constants of (16.1).

If certain modes, say p_μ, are coincident, the transient mesh solutions will be of the form

$$\begin{aligned} \{i\} &= J\{\gamma e^{pt}\}, \\ \{\gamma e^{pt}\} &= \{e^{p_1 t}, \cdots, e^{p_\mu t}, te^{p_\mu t}, \cdots, t^{q-1} e^{p_\mu t}, \cdots\}. \end{aligned} \tag{16.10}$$

The general solution $\{i\}_q$ of (14.2) is the sum of the steady-state and transient solutions. The charges can be found from $\{q\} = p^{-1}\{i\}$. The actual currents are given by Real $\{i\}_q$.

The use of elementary divisors in the transient solution of (14.1) is well known in the theory of differential equations.[27] This method will not be discussed here.

17. **Energy Relationships.**[28] The rate at which energy is being supplied to the network (14.1), the *instantaneous power*, is

$$(i)\{e\} = (i)A\{i\} = (i)R\{i\} + p(i)L\{i\}/2 + p(q)D\{q\}/2. \tag{17.1}$$

[26] See references (3), p. 168; (8).
[27] See references (4), (39), (58).
[28] See references (17), (19).

The rate at which energy is being converted into heat, the *total instantaneous power loss*, is

$$(17.2) \qquad \mathfrak{R} = (i)R\{i\} = \sum_{r,s=1}^{n} R_{rs}\, i_r\, i_s\,, \qquad R = (R_{rs})\,;$$

the rate of increase of magnetic energy, the *total instantaneous magnetic energy*, is

$$(17.3) \qquad \mathfrak{L} = (i)L\{i\}/2\,, \qquad L = (L_{rs})\,;$$

and the rate of increase of electric energy, the *total instantaneous electrostatic energy*, is

$$(17.4) \qquad \mathfrak{D} = (q)D\{q\}/2\,, \qquad D = (D_{rs})\,.$$

Thus, the theory of linear networks revolves to a considerable extent about the study of the three quadratic forms $\mathfrak{R}$, $\mathfrak{L}$, $\mathfrak{D}$, or their matrices R, L, D. (See §8.) The physical interpretations of the resistance (R) inductance (L) and elastance (D) matrices are evident.

18. **Equivalent Networks.**[29] Consider the linear network (14.1) with steady-state solution $\{I\} = B^{-1}\{E\}$. Let $\{I\}_m$ and $\{E\}_m$ denote $\{I\}$ and $\{E\}$ respectively, with all but the first m rows deleted, and $Y \equiv (Y_{st})$ denote B^{-1} with all but the first m rows and columns deleted. Then

$$(18.1) \qquad \{I\}_m = Y\{E\}_m\,.$$

Y is known as a *characteristic coefficient* (*admittance*) *matrix* of network A.

Two $2m$-pole linear networks are said to be *equivalent* if for all frequencies ($\omega = \lambda i$), they have equal characteristic coefficient matrices Y; i.e., for all ω, they have equal electrical characteristics.

Let the instantaneous power

$$(18.2) \qquad \mathfrak{B} = \mathfrak{R} + \mathfrak{L}\lambda + \mathfrak{D}\lambda^{-1} = (i)B\{i\} = (i)\{\epsilon\}$$

be subjected to the cogredient real $2m$-affine non-singular linear transformations (see §9)

$$(18.3) \qquad \{i\} = T\{\bar{\imath}\}\,, \qquad \{q\} = T\{\bar{q}\}\,.$$

Then $\mathfrak{B}$ becomes

$$(18.4) \qquad (\bar{\imath})\bar{B}\{\bar{\imath}\} = (\bar{\imath})\{\epsilon\}\,,$$

[29] See references (7), (9), (17), (19), (28).

where

$$\bar{B} = T^T \cdot B \cdot T . \tag{18.5}$$

Transformation T preserves the currents and charges in meshes $1, \cdots, m$. $\bar{B}$ may be considered to be a network matrix of a new network, structurally different from B, but having the same currents (and charges) in their meshes $1, \cdots, m$. If $\{\epsilon\}$ is invariant, $\mathfrak{B} = (i)\{\epsilon\} = (\bar{\imath})\{\bar{\epsilon}\}$ is an absolute invariant of B under T.

Since

$$\mathfrak{B} = (\bar{\imath})T^T \cdot B \cdot T\{\bar{\imath}\} = (\bar{\imath})T^T\{\epsilon\} = (\bar{\imath})\{\bar{\epsilon}\} ,$$

then

$$T^T\{\epsilon\} = \{\bar{\epsilon}\} ,$$

whence

$$\{\epsilon\} = (T^T)^{-1}\{\bar{\epsilon}\} \tag{18.6}$$

if the inverse of T exists. Thus, the currents (18.3) and voltages (18.6) are transformed contragrediently[30] under congruence (18.5).

The elements of Y are absolute invariants[30] of $\mathfrak{B}$ (and B) under T,

$$Y_{st} = d(B_t^s)/d(B) \qquad (s, t = 1, \cdots, m) \tag{18.7}$$

where B_t^s is B with row t and column s deleted, hence Y is an absolutely invariant matrix of B under T. Y_{st}, $s \neq t$, is the short-circuit transfer admittance between terminal pairs s and t, and Y_{ss} is the short-circuit driving-point admittance at terminal pairs s. Hence the networks B and $\bar{B}$ are equivalent, for Y is the characteristic coefficient matrix of both B and $\bar{B}$.

The ranks of B, B_t^s, Y are integer invariants. The rank of B^{-1} is the number of linearly independent mesh currents, the rank of Y the number of linearly independent driving-point currents.

If $I_1, \cdots, I_m$ are independent, the rank of Y is m and (18.1) gives

$$\{E\}_m = Z\{I\}_m \qquad Z \equiv (z_{st}) \equiv Y^{-1} . \tag{18.8}$$

Z, also known as a *characteristic coefficient* (*impedance*) *matrix* of network A, is an absolutely invariant matrix of B under T. z_{st}, $s \neq t$ is the open circuit transfer impedance between terminal pairs s and t, and z_{ss} the open circuit driving-point impedance between terminal pairs s.

[30] See references (9); (17), pp. 164–174.

The rank of Z is the number of linearly independent e.m.f's imposed across the terminal pairs $1, \cdots, m$.

Associated with A is a set of invariant matrices (§9) $A\begin{smallmatrix} s_1 & \cdots & s_t \\ r_1 & \cdots & r_t \end{smallmatrix}$; $(r_i, s_i \leqq m)$, matrices derived from A by deleting rows $r_1, \cdots, r_t$ and columns $s_1, \cdots, s_t$. The ranks, signatures and determinants of these matrices are invariants of A under congruence (18.5). These invariants play a significant part in the m-affine congruence theory. An extensive treatment of this subject, in which the mathematical structure underlying the theory of equivalent networks is developed, may be found in a forthcoming paper.[31]

Two networks may be equivalent and yet one or both of them physically not realizable. Necessary and sufficient conditions for the physical realizability of a network containing two types of circuit parameters have been discussed by Cauer[32] and others.

Thus it is seen how the equality relationships of m-affine congruence enters the theory, in the study of equivalent, though structurally different, networks.

19. **Symmetric Components.**[33] In the method of symmetric components, transformations P and Q of (14.2) are cogredient, and

$$P = Q = (p_{rs}), \qquad p_{rs} = a^{(r-1)(s-1)} \tag{19.1}$$

where a is a primitive n^{th} root of unity. Then $A\{i\} = \{\epsilon\}$ becomes

$$B\{\bar{\imath}\} = \{\bar{\epsilon}\}, \qquad B = P^{-1}AP, \tag{19.2}$$

where B is the network (or impedance) matrix in the new system of coordinates. The rows of P are known as *sequence operators* in the theory.

Thus the equality relationship of similarity enters the theory. In contrast with the case of the equality relationship of congruence—equivalent networks being structurally different—similarity involves only a change in reference system, the actual currents and voltages being unchanged.

20. **Direct Sum in Network Theory.** Consider two networks of (impedance) matrices A and B, respectively, and of orders α and β. Then (see §11) the direct sum $A + B$, of order $\alpha + \beta$, may be thought of as an (impedance) matrix of a new network formed by connecting networks A and B in such a manner as to form a net of $\alpha + \beta$ meshes and

[31] See reference (9).
[32] See references (13)–(19).
[33] See references (35), Vol. I, pp. 32–40.

so as to introduce no new mutual impedances or effect no changes in the total impedances a_{ij} and b_{ij}.

If the connection of A and B creates a linear dependence relation among the mesh currents, then an impedance matrix Q exists for the connected network which has fewer rows and columns than $A \dotplus B$. If the network be so connected that the relation between the currents $\{i\}$ before connection and after connection $\{i'\}$ is given by

$$\{i\} = C\{i'\}, \tag{20.1}$$

the (impedance) matrix after connection is

$$Q = C^T \cdot (A \dotplus B) \cdot C. \tag{20.2}$$

Since $A \dotplus B$ is symmetric, a property preserved under a congruence, Q is symmetric. (See Theorem 5.1.)

Theorem 20.1. *If a finite number of networks of (impedance) matrices $A, B, \cdots, M$ be connected in such a manner that the relation between the old and the new currents is given by* (20.1), *then the (impedance) matrix of the resulting network is*

$$Q = C^T \cdot (A \dotplus B \dotplus \cdots \dotplus M) \cdot C,$$

provided no new impedances are introduced.

Similar statements can be made with respect to the admittance matrix Y.

If A and B are equivalent network (impedance) matrices and B is decomposable into the direct sum of a (finite) number of network (impedance) matrices $B = B_1 \dotplus \cdots \dotplus B_m$, then the analysis of B becomes that of the components B_i. The theories developed earlier carry through for networks $A \dotplus \cdots \dotplus M$ as is evident from the theorems of §11. This treatment should clarify the use of the "connection tensor" of Kron[34] and similar decompositions appearing in the writings of Quade,[35] Cauer[36] and others, where the direct sum should be used rather than the ordinary sum.

21. **Further Results.** Quade[37] has classified free oscillations in linear passive networks (networks in which R, L, D of §17 are positive definite) of two degrees of freedom by means of elementary divisors. The first division is made according to the degree of the g. c. d. of the

[34] See references (35), Vol. I, pp. 50–56, also Vol. II–IV; (34), p. 147.
[35] See reference (43), pp. 14–22.
[36] See references (19), pp. 4–5; (17), p. 170.
[37] See reference (43).

elements of the network matrix A (see §7); the second division is made according to the types of roots the determinantal equation $D(\lambda) = 0$, possesses (see §16). The singular cases (where one or more of R, L, D are singular) is also considered. In this connection, Theorem 8.11 could have been used in a manner similar to that given in a forthcoming paper.[38]

Cauer,[39] in a series of papers, has considered certain problems in the theory of linear networks in which matric methods have appeared. In his theory of ideal transformers, the notion of linear dependence (§5) plays an important rôle. (An *ideal transformer* is characterized by a matrix of form $Z = \lambda L$, where L is of rank 1. Hence $L_{st} = Ku_s\, u_t$, where K is a constant, and not all the (real) u_s are zero. From $Z\{I\}_m = \{E\}_m$, $E_1 : E_2 : \cdots : E_m = u_1 : u_2 : \cdots : u_m$. Hold $E_1, \cdots, E_m$ fixed, let $K \to 0$ so that $u_1 I_1 + \cdots + u_m I_m = 0$. An ideal transformer may be characterized by the last two equations).

Gewertz[40] has considered the synthesis of 4-terminal linear passive networks; that is, the design of networks having such properties that they respond in a prescribed manner to the applied impulses. Matrices whose elements are rational algebraic functions of a complex variable play a leading part in this work. An extended bibliography is given at the end of Gewertz's paper.

The use of partial and continued (finite Stieltjes) fractions of matrices[41] have appeared in the literature in connection with problems in the theory of networks.

Kron[42] has studied the applications of tensors to electrical engineering problems, his principal papers being concerned with the theory of rotating machinery. A careful study of the methods of the present paper will show the carry over into the studies of Kron, in which the theory of differential forms[43] plays an important rôle, the methods of matric theory being applicable. The theory of skew matrices[44] appears to be useful in these studies and in the theory of non-linear circuits.

Goffman[45] has considered the problem of the m-affine congruence of pairs of forms in the real field, his results being applicable to the theory of equivalent networks.

[38] See reference (8).
[39] See references (12)–(19).
[40] See reference (23).
[41] See reference (19), (22).
[42] See references (33), (34), (35).
[43] See references (36), (40), (55).
[44] See reference (38).
[45] See reference (24).

REFERENCES

(1) Ahrens, W., *Über die Gleichungssysteme einer Kirchhoffschen galvanischen Stromverzweigung.* Math. Annalen 49 (1897) S. 311.

(2) Baerwald, H. G., *Der Geltungsbereich der Strecker-Feldtkellerschen Matrixgleichungen von Vierpolssystemen.* E. N. T. 1932, Heft 1, Band 9, Seite 31.

(3) Bieberbach-Bauer, *Algebra.* 4 Aufl., Leipzig 1928.

(4) Bocher, M., *Introduction to higher algebra.* MacMillan Co., N. Y., 1929.

(5) Brune, O., *Synthesis of a finite two-terminal network whose driving-point impedance is a prescribed function of frequency.* Sc. D. Thesis, M. I. T. 1931; J. of Math. and Phys. Vol. X, No. 3, (1931) p. 191.

(6) Burington, Richard S., *A classification of quadrics in affine n-space by means of arithmetic invariants.* Amer. Math. Monthly, Vol. 39 (1932), pp. 529–532.

(7) Burington, Richard S., *Invariants of quadrics and electric circuit theory.* Physical Review, Vol. 45 (1934), p. 429.

(8) Burington, Richard S., *A classification of plane cubic curves under the affine group by means of arithmetic invariants.* Toyoku Mathematical Journal, 1935.

(9) Burington, Richard S., *On the equivalence of quadrics in m-affine n-space and its relation to the equivalence of 2m-pole networks.* Transactions of Amer. Math. Soc., July 1935.

(10) Bush, V., *Operational circuit analysis.* John Wiley and Sons, New York, 1929.

(11) Carson, J. R., *Electric circuit theory and operational calculus.* McGraw Hill Book Co., New York, 1926.

(12) Cauer, W., *Die Verwirklichung von Wechselstrom-widerstanden vorgeschriebener Frequenzabhangigkeit.* Arch. f. Elektrotechn., B. 17, (1926-7) S. 355–388.

(13) Cauer, W., *Vierpole.* E. N. T., Band 6, Heft 7, (1929) S. 272.

(14) Cauer, W., *Über die Variablen eines passiven Vierpoles.* Sitzungsberichte der Preuss. Akademy der Wissenschaften. Dec. 1927.

(15) Cauer, W., *Untersuchungen über ein Problem, das drei positiv definite quadratische Formen mit Strecken-komplex in Bezichung setzt.* Math. Ann., Band 105, (1931) Seite 86.

(16) Cauer, W., *Siebschaltungen.* V. D. I., Verlag G. M. B. H., 1931.

(17) Cauer, W., *Ideal Transformatoren und lineare Transformationen.* E. N. T. 9 (1932), S. 157.

(18) Cauer, W., *Über Funktionen mit positivem Realteil.* Math. Annalen, Bd. 106, (1932), S. 369.

(19) Cauer, W., *Aquivalenz von 2n-polen ohne Ohmsche Widerstande.* Ges. d. Wiss. Nachrichten. Math.-Phys. Klasse. Fachgr. I. N. F. Bd. 1, (1934).

(20) Dickson, L. E., *Modern algebraic theories.* Chicago 1926.

(21) Foster, R. M., *A reactance theorem.* B. S. T. J., Vol. 3 (1924) p. 259.

(22) Fry, T. E., *The use of continued fractions in the design of electric networks.* Bull. of Amer. Math. Soc., Vol. XXV, p. 463.

(23) Gewertz, C. M., *Synthesis of a finite, four-terminal network from its pre-*

scribed driving-point functions and transfer function. J. of Math. and Phys., M. I. T., Vol. XII, (1933) pp. 1-256.

(24) GOFFMAN, C., *On the m-affine congruence of pairs of matrices and its relation to linear electrical networks.* B. S. Thesis, Case School of Applied Science, June 1935.

(25) GUILLEMIN, E. A., *Communication Networks.* Vol. I. (1931) John Wiley and Sons, New York.

(26) HEAVISIDE, O., *Electromagnetic theory.* Vol. 1, 2 and 3. London.

(27) HOLZER, H., *Die Beseitigung der Resonanzgefahr.* Schweiz. Bauzeitg. 82 (1923), S. 310.

(28) HOWITT, N., *Group theory and the electric circuit.* Physical Review, Vol. 37, (1931) pp. 1583-1595.

(29) HURWITZ, A., *Über die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen realen Teilen Besitzt.* Math. Annalen, 46, (1895) S. 273.

(30) JOHNSON, K. S., *Transmission circuits for telephonic communication.* Van Nostrand, New York (1927).

(31) KIRCHHOFF, G., *Über der Auflosung der Gleichungen auf welchen man bei der Untersuchung der linearen Verteilung galvanischer Strome gefuhrt wird.* Poggendorff s Annalen der Physik, Band 72, Ser. III-12 (1847) S. 497. Gesammelte Abhandlungen, Leipzig (1882).

(32) KOWALEWSKI, *Vorlesungen Über die Theorie der Determinanten.* Teubner (1909).

(33) KRON, G., *Generalized theory of electrical machinery.* Trans. A. I. E. E., (June 1930).

(34) KRON, G., *Non-Riemannian dynamics of rotating electrical machinery.* J. of Math. and Phys., M. I. T., Vol. XIII, No. 2 (1934) pp. 103-194.

(35) KRON, G., *The application of tensors to electrical engineering problems.* Vols. I-IV, Gen. Elect. Co. (1934). Gen. Elect. Review (1935).

(36) LEVI-CIVITA, *The absolute differential calculus.*

(37) MACDUFFEE, C. C., *The nullity of a matrix relative to a field.* Annals of Math., II, Vol. 27, No. 2, (1925), pp. 133-139.

(38) MACDUFFEE, C. C., *Matrices with elements in a principal ideal ring.* Bull. Amer. Math. Soc., Vol. 39 (1933), pp. 564-584.

(39) MACDUFFEE, C. C., *The theory of matrices.* Ergebnisse der Math., Berlin, 1933.

(40) MCCONNELL, *Applications of the absolute differential calculus.* Blackie and Sons.

(41) MOULTON, F. R., *Differential Equations.* MacMillan, 1930.

(42) ORE, O., *Some recent developments in abstract algebra.* Bull. Amer. Math. Soc., Vol. 37 (1931), pp. 537-548.

(43) QUADE, W., *Klassifikation der Schwingungsvorgange in gekoppelten Stromkreisen.* Universitatsverlag von Robert Noske in Borna-Leipzig, 1933.

(44) RAYLEIGH, *The reaction upon the driving point of a system ezecuting forced harmonic oscillations of various periods.* Scientific Papers, Vol. II, 1881-1887, Cambridge (1900).

(45) ROUTH, E. J., *Advanced Rigid Dynamics.* MacMillan and Co., Lt. (1905).

(46) SELACH, E., *Zur Theorie der Vierpolverbindungen.* E. N. T., 1931, Bd. 8, Heft 7, S. 297.

(47) SHEA, T. E., *Transmission networks and wave filters.* Van Nostrand, New York, 1929.

(48) STRECKER, F. UND FELDTKELLER, R., *Grundlagen der Theorie des allgemeinen Vierpols.* E. N. T., 1929, Bd. 6, Heft 3, S. 93.

(49) THEVENIN, M. L., *Sur un Nouveau Theorem d Electricite Dynamique.* Comptes Rendus, Tome 97, (1883) p. 159.

(50) THOMSON, J. J., *Recent researches in electricity and magnetism.* Chapt. VI: *The distribution of rapidly alternating currents.* Oxford, 1893.

(51) TROTT, G. R., *On the canonical form of a non-singular pencil of Hermitian matrices.* Amer. J. of Math., Vol. LVI, (1934) pp. 359–371.

(52) TURNBULL, *The theory of determinants, matrices, invariants.* 1928, Blackie and Son.

(53) TURNBULL AND AITKEN, *Canonical matrices.* Glasgow, 1932.

(54) VAN DER WAERDEN, *Moderne Algebra.* Vols. I and II, Berlin, J. Springer 1930.

(55) VEBLEN, *Invariants of quadratic differential forms.* Cambridge Tracts in Math., No. 24.

(56) VON HELMHOLTZ, H., *Über die Dauer und den Verlauf der durch Stromesschwankungen induzierten elektrischen Strome.* Poggendorfs Annalen 83 (1851), S. 505.

(57) WEDDERBURN, J. H. M., *Lectures on matrices.* Amer. Math. Soc., Colloq. Pub., Vol. XVII (1934).

(58) WHITTAKER, *Analytic dynamics.* Cambridge, 1917.

ALMOST PERIODIC FUNCTIONS AND THE VIBRATING MEMBRANE

By Arra Steve Avakian

Introduction

The purpose of this paper is to illustrate the methods of almost periodic functions of Harald Bohr[1] in the solution of certain wave equations.

The specific problem treated will be that of the circular vibrating membrane with non-uniform mass and tension. Under certain conditions the displacement is an almost periodic function of time.

The method employs the finiteness of the energy of the system. This was first employed by C. Muckenhoupt,[2] and considered under more general conditions by S. Bochner.[3]

The bibliography at the end of this paper presents only a representative list of works in the field of almost periodic functions.

The author of this thesis gratefully acknowledges the guidance and assistance of Professor Philip Franklin in the preparation of this work.

I. Theory of Functions Almost Periodic in the Mean

We will need to develop certain theorems on functions almost periodic in the mean, and for later application to the problem of the vibrating membrane we shall use two space parameters x and y. It is apparent that the theorems hold also for the space parameters r and θ, in polar coördinates.

Our variables will be assumed real, but functions may take on complex values. This will permit the use of exponentials instead of trigonometric functions.

We may write

$$f(t; x, y) = R(t; x, y) + iI(t; x, y)$$

where the real part $R(t; x, y)$ only will appear in physical problems.

[1] See bibliography 5.
[2] See bibliography 8.
[3] See bibliography 4.

Let the function $f(t; x, y)$ be defined in a region R

$$(x_1 \leqq x \leqq x_2; y_1 \leqq y \leqq y_2; -\infty < t < +\infty);$$

quadratically summable in x and y for all t; continuous in t; and continuous in the mean in x and y, that is, for any given ϵ there exists a δ such that

$$\int\int_R |f(t; x+\delta, y+\delta) - f(t; x, y)|^2 \, dxdy < \epsilon$$

for all t.

Displacement numbers τ will be taken in the direction of the t axis; these will be defined below.

We say τ is a *displacement number* of $f(t; x, y)$ belonging to ϵ if

$$\int\int_R |f(t+\tau; x, y) - f(t; x, y)|^2 \, dxdy \leqq \epsilon$$

uniformly for all t.

A function $f(t; x, y)$ is said to be *almost periodic in the mean* in t in any region as above if all the possible displacement numbers of $f(t; x, y)$ belonging to any given ϵ form a relatively dense set of numbers along the t axis. The minimum interval along the t axis which always contains at least one of these displacement numbers is called the *inclusion interval*, it depends on ϵ and it is written l_ϵ. When this interval is placed beginning at the origin it is called the initial interval.

The following theorems on the properties of functions almost periodic in the mean are valid for any number of space parameters since the proofs may readily be extended. We will omit those proofs which may be found in the paper by C. Muckenhoupt, already mentioned.

THEOREM I. Every function almost periodic in the mean is bounded in the mean.

COROLLARY 1. *If $f(t; x, y)$ is a function almost periodic the mean, then the square $\{f(t; x, y)\}^2$ is also almost periodic in the mean.*

For, from Schwarz' inequality we may write

$$\int\int_R |\{f(t+\tau; x, y)\}^2 - \{f(t; x, y)\}^2|^2 \, dxdy$$

$$= \int\int_R [|f(t+\tau; x, y) - f(t; x, y)| \cdot |f(t+\tau; x, y) + f(t; x, y)|]^2 dxdy$$

$$\leqq \left[\int\int_R |f(t+\tau;x,y) - f(t;x,y)|^4\,dxdy\right]^{\frac{1}{4}} \cdot \left[\int\int_R |f(t+\tau;x,y) + f(t;x,y)|^4\,dxdy\right]^{\frac{1}{4}}$$

$$\leqq \left[\int\int_R |f(t+\tau;x,y) - f(t;x,y)|^2\,dxdy\right] \cdot \left[\int\int_R |f(t+\tau;x,y) + f(t;x,y)|^2\,dxdy\right]$$

$$\leqq \epsilon \cdot B \leqq \epsilon_1$$

since the inequality $\int\int \varphi^2 \leqq \left[\int\int \varphi\right]^2$ is true if φ is always positive.

COROLLARY 2. The absolute value of a function almost periodic in the mean is likewise almost periodic in the mean.

THEOREM II. Every function almost periodic in the mean is uniformly continuous in the mean.

COROLLARY. If τ_1 is a displacement number for $f(t;x,y)$ belonging to ϵ_1, then there exists a $\delta > 0$ depending on ϵ_1 and an $\epsilon_2 > \epsilon_1$ such that if $|\tau_2 - \tau_1| < \delta$ then τ_2 is a displacement number belonging to ϵ_2.

Let δ be chosen according to Theorem II such that for any two points t' and t'' where $|t' - t''| < \delta$ we have

$$\int\int_R |f(t';x,y) - f(t'';x,y)|^2\,dxdy \leqq \frac{\epsilon_2 - \epsilon_1}{2(1+k^2)}$$

where k is any real number large enough such that

$$\left(1 + \frac{1}{k^2}\right)\epsilon_1 < \frac{\epsilon_2 + \epsilon_1}{2}.$$

Then from the inequality

$$(a+b)^2 \leqq (a+b)^2 + \left(ka - \frac{b}{k}\right)^2 = a^2(1+k^2) + b^2\left(1 + \frac{1}{k^2}\right)$$

we have

$$\int\int_R |f(t+\tau_2;x,y) - f(t;x,y)|^2\,dxdy$$

$$= \int\int_R |f(t+\tau_2;x,y) - f(t+\tau_1;x,y) + f(t+\tau_1;x,y) + f(t;x,y)|^2\,dxdy$$

$$\leq (1 + k^2) \int\int_R |f(t + \tau_2; x, y) - f(t + \tau_1; x, y)|^2 \, dxdy$$
$$+ \left(1 + \frac{1}{k^2}\right) \int\int_R |f(t + \tau_1; x, y) - f(t; x, y)^2 \, dxdy$$
$$\leq \frac{\epsilon_2 - \epsilon_1}{2} + \frac{\epsilon_2 + \epsilon_1}{2} = \epsilon_2 .$$

THEOREM III. Any two functions almost periodic in the mean are simultaneously almost periodic in the mean; that is we may find an inclusion interval l_ϵ, such that in every interval l_ϵ there is at least one displacement number belonging to ϵ for both functions simultaneously.

COROLLARY. Any n functions almost periodic in the mean are simultaneously almost periodic in the mean; that is we may find an inclusion interval l_ϵ such that in every interval l_ϵ there is at least one displacement number belonging to ϵ for all n functions simultaneously.

THEOREM IV. The sum of n functions almost periodic in the mean is likewise almost periodic in the mean.

THEOREM V. The product of a function almost periodic in the mean with a function of one variable almost periodic in that variable is itself almost periodic in the mean.

COROLLARY. The product of a function almost periodic in the mean with n functions of one variable almost periodic in that variable is itself almost periodic in the mean.

THEOREM VI. Any uniformly convergent series of functions almost periodic in the mean is likewise almost periodic in the mean.

COROLLARY. If the series $\sum_{k=0}^{\infty} A_k(x, y)e^{iA_k t}$ converges uniformly for all values of x and y in R, and hence in particular if $\sum_{k=0}^{\infty} A_k(x, y)$ converges similarly, then it represents a function almost periodic in the mean in t.

THEOREM VII. If $f(t; x, y)$ is almost periodic in the mean, then

$$\text{l.i.m.}_{T \to \infty} \frac{1}{T} \int_A^{A+T} f(t; x, y) \, dt$$

exists and is independant of A.

This limit is called the mean value of $f(t; x, y)$ and it is abbreviated by $M\{f(t; x, y)\}$ and is a function of x and y alone.

Let l be the inclusion interval for $f(t; x, y)$ belonging to ϵ. Let the

first displacement number greater than l be denoted by τ_ϵ. Take an interval in t of length τ_ϵ beginning as at $t = A$ and ending at $t = A + \tau_\epsilon$. Now consider the integral.

$$I = \int\int_R \left| \frac{1}{\tau_\epsilon} \int_A^{A+\tau_\epsilon} f(t; x, y)\, dt - \frac{1}{\tau_\epsilon} \int_0^{\tau_\epsilon} f(t; x, y)\, dt \right|^2 dxdy .$$

Let T_ϵ be a displacement number belonging to ϵ and in the interval from A to $A + \tau_\epsilon$. T_ϵ must exist since $\tau_\epsilon > l$. Now split the above integral into the following parts:

$$I = \int\int_R \left| \frac{1}{\tau_\epsilon} \int_{T_\epsilon}^{A+\tau_\epsilon} f(t; x, y)\, dt - \frac{1}{\tau_\epsilon} \int_0^{A+\tau_\epsilon-T_\epsilon} + \frac{1}{\tau_\epsilon} \int_A^{T_\epsilon} - \frac{1}{\tau_\epsilon} \int_{A+\tau_\epsilon-T_\epsilon}^{\tau_\epsilon} \right|^2 dxdy .$$

After rewriting, using a translation in the variable t with a corresponding shift in the interval of integration, and grouping we obtain

$$I = \int\int_R \left| \frac{1}{\tau_\epsilon} \int_0^{A+\tau_\epsilon-T_\epsilon} [f(t + T_\epsilon; x, y) - f(t; x, y)]\, dt + \frac{1}{\tau_\epsilon} \int_{A+\tau_\epsilon-T_\epsilon}^{\tau_\epsilon} [f(t + T_\epsilon - \tau_\epsilon; x, y) - f(t; x, y)]\, dt \right|^2 dxdy$$

Since T_ϵ and τ_ϵ are displacement numbers belonging to ϵ then $T_\epsilon - \tau_\epsilon$ is a displacement number belonging to 4ϵ. Then by noting that

$$\int_0^{A+\tau_\epsilon-T_\epsilon} dt \leqq \tau_\epsilon \text{ and } \int_{A+\tau_\epsilon-T_\epsilon}^{\tau_\epsilon} dt \leqq \tau_\epsilon$$

and by using the Schwarz inequality we obtain the following chain of inequalities from the above integral.

$$I \leqq \frac{2}{\tau_\epsilon^2} \int\int_R \left| \int_0^{A+\tau_\epsilon-T_\epsilon} [f(t + T_\epsilon; x, y) - f(t; x, y)]\, dt \right|^2 dxdy + \frac{2}{\tau_\epsilon^2} \int\int_R \left| \int_{A+\tau_\epsilon-T_\epsilon}^{\tau_\epsilon} [f(t + T_\epsilon - \tau_\epsilon; x, y) - f(t; x, y)]\, dt \right|^2 dxdy$$

$$\leqq \frac{2}{\tau_\epsilon^2} \int\int_R \left\{ \int_0^{A+\tau_\epsilon-T_\epsilon} dt \cdot \int_0^{A+\tau_\epsilon-T_\epsilon} |f(t + T_\epsilon; x, y) - f(t; x, y)|^2\, dt \right\} dxdy$$

$$+\frac{2}{\tau_\epsilon^2}\iint_R\left\{\int_{A+\tau_\epsilon-T_\epsilon}^{\tau_\epsilon}dt\cdot\int_{A+\tau_\epsilon-T_\epsilon}^{\tau_\epsilon}|f(t+T_\epsilon-\tau_\epsilon;x,y)-f(t;x,y)|^2\,dt\right\}$$

$$dxdy$$

$$\leqq\frac{2}{\tau_\epsilon}\int_0^{A+\tau_\epsilon-T_\epsilon}\left\{\iint_R|f(t+T_\epsilon;x,y)-f(t;x,y)|^2\,dxdy\right\}dt$$

$$+\frac{2}{\tau_\epsilon}\int_{A+\tau_\epsilon-T_\epsilon}^{\tau_\epsilon}\left\{\iint_R|f(t+T_\epsilon-\tau_\epsilon;x,y)-f(t;x,y)|^2\,dxdy\right\}dt$$

$$\leqq\frac{2}{\tau_\epsilon}\int_0^{A+\tau_\epsilon-T_\epsilon}\epsilon\cdot dt+\frac{2}{\tau_\epsilon}\int_{A+\tau_\epsilon-T_\epsilon}^{\tau_\epsilon}4\epsilon\cdot dt\leqq 2\epsilon+8\epsilon=10\epsilon=\epsilon_1.$$

Thus we have shown that

$$I=\iint_R\left|\frac{1}{\tau_\epsilon}\int_A^{A+\tau_\epsilon}f(t;x,y)\,dt-\frac{1}{\tau_\epsilon}\int_0^{\tau_\epsilon}f(t;x,y)\,dt\right|^2dxdy\leqq\epsilon_1$$

for an arbitrarily chosen interval of length τ_ϵ along t.

Now for any interval along t we may divide it into a whole number of sub-intervals of length τ_ϵ (for each one of which the above inequality holds) and a remainder interval of length less than τ_ϵ. For the n intervals of length τ_ϵ we know

$$I'=\iint_R\left|\frac{1}{n\tau_\epsilon}\int_A^{A+n\tau_\epsilon}f(t;x,y)\,dt-\frac{1}{\tau_\epsilon}\int_0^{\tau_\epsilon}f(t;x,y)\,dt\right|^2dxdy\leqq\epsilon_2$$

by a proof similar to the above where if we write the coefficient of the second integral as $n/n\tau_\epsilon$ instead of $1/\tau_\epsilon$ we may by using the inequality $(\Sigma a)^2\leqq n\cdot\Sigma a^2$ separate I' into n times the sum of n integrals like I with limits $A+(k-1)\tau_\epsilon$ to $A+k\tau_\epsilon$ and coefficients $1/n\tau_\epsilon$. Adding these we get upon cancelling n^2 in both numerator and denominator an integral exactly as I.

We may now consider the remainder interval of length $\theta\tau_\epsilon$ where $0<\theta<1$. If we set up the integral

$$J=\iint_R\left|\frac{1}{n\tau_\epsilon+\theta\tau_\epsilon}\int_A^{A+n\tau_\epsilon+\theta\tau_\epsilon}f(t;x,y)\,dt\right.$$

$$\left.-\frac{1}{n\tau_\epsilon}\int_A^{A+n\tau_\epsilon}f(t;x,y)\,dt\right|^2dxdy$$

we may immediately rewrite it as $J =$

$$\iint_R \Big| - \frac{\theta\tau_\epsilon}{n\tau_\epsilon(n\tau_\epsilon + \theta\tau_\epsilon)} \int_A^{A+n\tau_\epsilon} f(t; x, y)\, dt + \frac{1}{n\tau_\epsilon + \theta\tau_\epsilon} \int_{A+n\tau_\epsilon}^{A+n\tau_\epsilon+\theta\tau_\epsilon} f(t; x, y)\, dt \Big|^2 dxdy$$

$$\leqq 2 \iint_R \Big| - \frac{\theta\tau_\epsilon}{n\tau_\epsilon(n\tau_\epsilon + \theta\tau_\epsilon)} \int_A^{A+n\tau_\epsilon} f(t; x, y)\, dt \Big|^2 dxdy + 2 \iint_R \Big| \frac{1}{n\tau_\epsilon + \theta\tau_\epsilon} \int_{A+n\tau_\epsilon}^{A+n\tau_\epsilon+\theta\tau_\epsilon} f(t; x, y)\, dt \Big|^2 dxdy$$

$$\leqq 2 \iint_R \left[\left\{ \frac{\theta\tau_\epsilon}{n\tau_\epsilon(n\tau_\epsilon + \theta\tau_\epsilon)} \right\}^2 \cdot \max_t |f(t; x, y)|^2 \left\{ \int_A^{A+n\tau_\epsilon} dt \right\}^2 \right] dxdy + 2 \iint_R \left[\frac{1}{(n\tau_\epsilon + \theta\tau_\epsilon)^2} \max_t |f(t; x, y)|^2 \left\{ \int_{A+n\tau_\epsilon}^{A+n\tau_\epsilon+\theta\tau_\epsilon} dt \right\}^2 \right] dxdy$$

$$= 2 \iint_R \left[\left(\frac{\theta\tau_\epsilon}{n\tau_\epsilon + \theta\tau_\epsilon} \right)^2 \max_t |f(t; x, y)|^2 \right] dxdy + 2 \iint_R \left[\left(\frac{\theta\tau_\epsilon}{n\tau_\epsilon + \theta\tau_\epsilon} \right)^2 \max_t |f(t; x, y)|^2 \right] dxdy$$

$$= 4 \left(\frac{\theta\tau_\epsilon}{n\tau_\epsilon + \theta\tau_\epsilon} \right)^2 \iint_R \max_t |f(t; x, y)|^2\, dxdy .$$

Since $f(t; x, y)$ is bounded in t and bounded in the mean in x and y, the above integral and therefore the integral J approaches zero as n approaches infinity.

Therefore, if we write any interval $B - A$ as $n\tau_\epsilon + \theta\tau_\epsilon$, we have

$$\lim_{B-A\to\infty} \iint_R \Big| \frac{1}{B-A} \int_A^B f(t; x, y)\, dt - \frac{1}{\tau_\epsilon} \int_0^{\tau_\epsilon} f(t; x, y)\, dt \Big|^2 dxdy$$

$$\leqq \lim_{n\to\infty} 2 \iint_R \Big| \frac{1}{n\tau_\epsilon + \theta\tau_\epsilon} \int_A^{A+n\tau_\epsilon+\theta\tau_\epsilon} f(t; x, y)\, dt - \frac{1}{n\tau_\epsilon} \int_A^{A+n\tau_\epsilon} f(t; x, y)\, dt \Big|^2 dxdy$$

$$+ \lim_{n\to\infty} 2 \iint_R \Big| \frac{1}{n\tau_\epsilon} \int_A^{A+n\tau_\epsilon} f(t; x, y)\, dt - \frac{1}{\tau_\epsilon} \int_0^{\tau_\epsilon} f(t; x, y)\, dt \Big|^2 dxdy .$$

The first limit is $\lim_{n\to\infty} J = 0$. The second integral is I' which is uniformly less than ϵ_2 and which therefore has a limit less than ϵ_2. It follows then that

$$\text{l.i.m.}_{B-A\to\infty} \frac{1}{B-A} \int_A^B f(t; x, y)\, dt$$

exists where A has been held fixed. Now A has been arbitrary throughout the proof. Therefore, if we write $T = B - A$, we have shown that

$$\text{l.i.m.}_{T\to\infty} \frac{1}{T} \int_A^{A+T} f(t; x, y)\, dt$$

exists and is independant of A, and the theorem is proved.

We are now ready to consider approximations to functions almost periodic in the mean by trigonometric (exponential) polynomials. We observe first of all for any real λ, $e^{i\lambda t}$ is a purely periodic and therefore a fortiori an almost periodic function. If now $f(t; x, y)$ is a function almost periodic in the mean in t, then, according to Theorem V, $f(t; x, y)\cdot e^{-i\lambda t}$ is a function almost periodic in the mean. Then, according to Theorem VII, the mean value exists and is

$$M\{f(t; x, y)e^{-i\lambda t}\} = a(\lambda; x, y)\,.$$

The proofs of the following four theorems follow those of A. Besicovitch.[4]

THEOREM VIII. If $f(t; x, y)$ is a function almost periodic in the mean; $\lambda_1, \lambda_2, \cdots, \lambda_N$ N arbitrary real numbers mutually different; and $b_1, b_2, \cdots, b_N$ N arbitrary real or complex numbers; then

$$M\{|\, f(t; x, y) - \sum_1^N b_n e^{i\lambda_n t}|^2\}$$

$$= M\{|\, f(t; x, y)\,|^2\} - \sum_1^N |\, a(\lambda_n; x, y)\,|^2 + \sum_1^N |\, b_n - a(\lambda_n; x, y)\,|^2\,.$$

This is called the *equation of approximation in the mean.* Here we see that for fixed λ_n the polynomial $\sum_1^N b_n e^{i\lambda_n t}$ gives the best approximation to $f(t; x, y)$ if we take $b_n = a(\lambda_n; x, y)$. Then since the left side of the equation is non-negative we have

$$\sum_1^N |\, a(\lambda_n; x, y)\,|^2 \leqq M\{|\, f(t; x, y)\,|^2\}\,.$$

[4] See bibliography 1.

THEOREM IX. There exist at most an enumerably infinite set of values λ *for which* $a(\lambda; x, y)$ *differs from zero.*

We denote these values of λ by $\Lambda_1, \Lambda_2, \cdots$ and set $a(\Lambda_n; x, y) = A_n$ for all n. The nunbers A_n are the Fourier coefficients of the function $f(t; x, y)$. We may write

$$f(t; x, y) \sim \sum_1^\infty A_n e^{i\Lambda_n t},$$

then from Theorem VIII we have

$$\sum_1^\infty |A_n|^2 \leq M\{|f(t; x, y)|^2\}.$$

This is the *Bessel inequality* for functions almost periodic in the mean. In the case of purely periodic functions this definition of Fourier series corresponds to the ordinary one.

THEOREM X. The Fourier series of a function almost periodic in the mean represented by a uniformly convergent trigonometric series

$$f(t; x, y) = \sum_1^\infty a_n(x, y) e^{i\lambda_n t}$$

is coincident with this series.

Here we find that

$$M\{f(t; x, y)e^{-i\lambda_n t}\} = a_n(x, y)$$

for all n which proves the theorem.

THEOREM XI. BOHR'S FUNDAMENTAL THEOREM. If A_n *denote the non-vanishing Fourier coefficients of* $f(t; x, y)$ *almost periodic in the mean, then*

$$\sum_1^\infty |A_n|^2 = M\{|f(t; x, y)|^2\}.$$

This theorem states that, better than the Bessel inequality, the *Parseval equation* holds.

We are now ready to prove Bohr's approximation theorem. This theorem was first proved by H. Bohr,[5] later by S. Bochner,[6] and still later by H. Weyl[7] who used the Parseval equation. We shall here give Weyl's proof according to Besicovitch.

[6] See bibliography 5.
[6] See bibliography 10.
[7] See bibliography 9.

THEOREM XII. APPROXIMATION THEOREM. If there is given a function almost periodic in the mean

$$f(t; x, y) \sim \sum_1^\infty A_n(x, y)e^{iA_n t}$$

and an $\epsilon > 0$, then there exists a trigonometric polynomial $P_\epsilon(t; x, y)$ whose exponents are the Fourier exponents of $f(t; x, y)$ above, and which satisfies the inequality

$$\int\int_R |f(t; x, y) - P_\epsilon(t; x, y)|^2 \, dxdy \leqq \epsilon$$

for all values of t.

We write

$$f(t; x, y) - \sum_1^p A_n(x, y)e^{iA_n t} = f_p(t; x, y),$$

and by the Parseval equation

$$M\{|f_p(t; x, y)|^2\} = \sum_{n=p+1}^\infty |A_n|^2.$$

Then given any $\eta > 0$ it is possible to find a p such that

$$M\{|f_p(t; x, y)|^2\} < \eta \tag{1}$$

since the series $\Sigma |A_n|^2$ is convergent. Furthermore, since $f_p(t; x, y)$ is almost periodic in the mean, $|f_p(t; x, y)|^2$ is also almost periodic in the mean. Then from the theorem of the mean value, since

$$\frac{1}{T}\int_A^{A+T} |f_p(t; x, y)|^2 \, dt \to M\{|f_p(t; x, y)|^2\}$$

for any A as $P \to \infty$, there exists a $T_0 > 0$ such that

$$\left| \frac{1}{T}\int_0^T |f_p(t + s; x, y)|^2 \, dt - M\{|f_p(t; x, y)|^2\} \right| \leqq \eta \tag{2}$$

for all $T > T_0$ and for all s. Then from the inequalities (1) and (2) we have

$$\frac{1}{T}\int_0^T |f_p(t + s; x, y)|^2 \, dt < 2\eta. \tag{3}$$

Let l be the inclusion interval belonging to $\epsilon/9$. Take a

$$T = N(l + 1) > T_0$$

where N is an integer, and in each interval

$$[k(l+1), \quad k(l+1)+l] \qquad (k = 0, 1, \cdots, N-1)$$

of length l along the t axis take a τ_k belonging to $\epsilon/9$. Now choose a δ where $0 < \delta < 1$ such that

$$\int\int_R |f(t''; x, y) - f(t'; x, y)|^2 \, dxdy \leqq \epsilon/9$$

for all t'' and t' if only $|t'' - t'| < \delta$. We define $e(t)$ in the interval $[0, T]$ as follows:

$e(t) = 1$ in all intervals $[\tau_k, \tau_k + \delta]$, $k = 0, 1, \cdots, N-1$,

$e(t) = 0$ at all other points.

The intervals $[\tau_k, \tau_k + \delta]$ do not entirely overlap since they are of length $\delta < 1$ and they have their lower end point in intervals which start at points greater than one unit apart. From the Schwarz inequality we have

$$\left| \int_0^T f_p(t+s; x, y) e(t) \, dt \right|^2 \leqq \int_0^T |f_p(t+s; x, y)|^2 \, dt \cdot \int_0^T (e(t))^2 \, dt \quad (4)$$

Since

$$\int_0^T f_p(t+s; x, y) e(t) \, dt = \sum_{k=0}^{N-1} \int_{\tau_k}^{\tau_k+\delta} f_p(t+s; x, y) \, dt$$

$$= \sum_{k=0}^{N-1} \int_0^{\delta} f_p(t+\tau_k+s; x, y) \, dt$$

and since

$$\int_0^T (e(t))^2 \, dt = N\delta,$$

then from (3) and (4) we have

$$\left| \sum_{k=0}^{N-1} \int_0^{\delta} f_p(t+\tau_k+s; x, y) \, dt \right|^2 < 2TN\eta\delta$$

or

$$\left| \frac{1}{N\delta} \sum_{k=0}^{N-1} \int_0^{\delta} f_p(t+\tau_k+s; x, y) \, dt \right|^2 < \frac{2T\eta}{N\delta} = \frac{2\eta(l+1)}{\delta}$$

Now if we choose

$$\eta < \frac{\epsilon\delta}{18(l+1)S}$$

where S is the area of the region of integration R, we have

$$\iint_R \left| \frac{1}{N\delta} \sum_{k=0}^{N-1} \int_0^\delta f_p(t+\tau_k+s;x,y)\,dt \right|^2 dxdy \leqq \frac{2\eta(l+1)}{\delta} \iint_R dxdy = \frac{2\eta(l+1)}{\delta} S < \epsilon/9. \tag{5}$$

But

$$\frac{1}{N\delta}\int_0^\delta f_p(t+\tau_k+s;x,y)\,dt = \frac{1}{N\delta}\int_0^\delta f(t+\tau_k+s;x,y)\,dt - P_k(s;x,y)$$

where

$$P_k(s;x,y) = \frac{1}{N\delta}\int_0^\delta \sum_{n=1}^{p} A_n(x,y)e^{iA_n(t+\tau_k+s)}\,dt = \frac{1}{N\delta}\sum_{n=1}^{p} e^{iA_n s}\int_0^\delta A_n(x,y)e^{iA_n(t+\tau_k)}\,dt.$$

$P_k(s;x,y)$ is a trigonometric polynomial with exponents belonging to the Fourier exponents of $f(t;x,y)$. We may now rewrite (5) as

$$\iint_R \left| \frac{1}{N\delta} \sum_{k=0}^{N-1} \int_0^\delta f(t+\tau_k+s;x,y)\,dt - P(s;x,y) \right|^2 dxdy < \epsilon/9 \tag{6}$$

where

$$P(s;x,y) = \sum_{k=0}^{N-1} P_k(s;x,y).$$

But we have previously chosen δ such that

$$\iint_R |f(t+\tau_k+s;x,y) - f(\tau_k+s;x,y)|^2\,dxdy < \epsilon/9$$

for $0 \leqq t \leqq \delta$. Consequently,

$$\iint_R \left| \frac{1}{\delta}\int_0^\delta f(t+\tau_k+s;x,y)\,dt - f(\tau_k+s;x,y) \right|^2 dxdy < \epsilon/9. \tag{7}$$

But since

$$\sum_{k=0}^{N-1}\left\{\frac{1}{\delta}\int_0^{\delta} f(t+\tau_k+s;\,x,\,y)\,dt - f(\tau_k+s;\,x,\,y)\right\}$$

$$\geqq \left|\frac{1}{\delta}\sum_{k=0}^{N-1}\int_0^{\delta} f(t+\tau_k+s;\,x,\,y)\,dt - Nf(\tau_{k_1}+s;\,x,\,y)\right|$$

for some value of k as k_1, the inequality (7) becomes upon summing on k

$$\int\int_R \left|\frac{1}{N\delta}\sum_{k=0}^{N-1}\int_0^{\delta} f(t+\tau_k+s;\,x,\,y)\,dt - f(\tau_{k_1}+s;\,x,\,y)\right|^2 dxdy < \epsilon/9\,. \quad (8)$$

Now since $f(t;\,x,\,y)$ is almost periodic in the mean we have

$$\int\int_R |\,f(\tau_{k_1}+s;\,x,\,y) - f(s;\,x,\,y)\,|^2\,dxdy \leqq \epsilon/9\,. \quad (9)$$

Then by combining (6), (8), and (9) we obtain

$$\int\int_R |f\,(s;\,x,\,y) - P(s;\,x,\,y)\,|^2\,dxdy \leqq 3(\epsilon/9+\epsilon/9+\epsilon/9) = \epsilon\,,$$

and the theorem on the approximation is proved.

Hence from this, Bohr's Approximation Theorem, we may approximate to any degree of accuracy any function almost periodic in the mean by a trigonometric polynomial.

II. Proof of the Almost Periodicity in the Mean of the Solution of the Vibrating Membrane

We shall now consider a vibrating circular membrane whose mass and tension are functions of position, but not of time. We shall use polar and rectangular coördinates. Let the mass per unit area be $\mu(r,\,\theta)$; let the tension in the direction of r be $T_1(r,\,\theta)$, and the tension in the direction of θ be $T_2(r,\,\theta)$.

From $F = ma$ we may set up the differential equation for the displacement of the membrane $z = z(t;\,r,\,\theta)$ by taking vertical components of the tension on the four sides of an element of area $dA = rd\theta dr$.

We obtain

$$(r + dr)\, d\theta\, T_1(r + dr, \theta) \frac{\partial z}{\partial r}\bigg|_{r+dr} - rd\theta\, T_1(r, \theta) \frac{\partial z}{\partial r}\bigg|_{r}$$

$$+ dr\, T_2(r, \theta + d\theta) \frac{\partial z}{\partial s}\bigg|_{\theta+d\theta} - dr T_2(r, \theta) \frac{\partial z}{\partial s}\bigg|_{\theta} = \mu(r, \theta)\, rdrd\theta \frac{\partial^2 z}{\partial t^2}$$

where $\frac{\partial}{\partial s} = \frac{1}{r} \frac{\partial}{\partial \theta}$

This becomes

$$dr \frac{\partial}{\partial r}\left[rd\theta T_1(r, \theta) \frac{\partial z}{\partial r}\right] + rd\theta \frac{\partial}{\partial s}\left[dr T_2(r, \theta) \frac{\partial z}{\partial s}\right] = \mu(r, \theta)\, rdrd\theta \frac{\partial^2 z}{\partial t^2}$$

or

$$\frac{1}{r} \frac{\partial}{\partial r}\left[rT_1(r, \theta) \frac{\partial z}{\partial r}\right] + \frac{1}{r^2} \frac{\partial}{\partial \theta}\left[T_2(r, \theta) \frac{\partial z}{\partial \theta}\right] = \mu(r, \theta) \frac{\partial^2 z}{\partial t^2} \qquad (1)$$

which is the differential equation in polar coördinates for the displacement. For stability in the vibrating system the functions μ, T_1, and T_2 must be chosen always positive and bounded. They may be taken differentiable as often as will be necessary.

If the tension is uniform, then $T_1 = T_2 = T$ = constant and (1) reduces to the wave equation in polar coördinates

$$\frac{1}{r} \frac{\partial}{\partial r}\left(r \frac{\partial z}{\partial r}\right) + \frac{1}{r^2} \frac{\partial^2 z}{\partial \theta^2} = \frac{\mu}{T} \frac{\partial^2 z}{\partial t^2}$$

In rectangular coördinates the differential equation is

$$\frac{\partial}{\partial x}\left[T_3(x, y) \frac{\partial z}{\partial x}\right] + \frac{\partial}{\partial y}\left[T_4(x, y) \frac{\partial z}{\partial y}\right] = \mu_1(x, y) \frac{\partial^2 z}{\partial t^2} \qquad (1')$$

where T_3, T_4, and μ_1 are necessarily different from T_1, T_2, and μ, but having the same properties.

We remark here that the equation (1) in polar coördinates does not come under the general case in n dimensions discussed by S. Bochner[4] in his paper, but the equation (1′) in rectangular coördinates does.

We may set up the kinetic energy $1/2\ mv^2$ of the membrane as

$$\frac{1}{2} \int_0^R \int_0^{2\pi} \mu(r, \theta) \left(\frac{\partial z}{\partial t}\right)^2 rdrd\theta$$

[4] See bibliography 4.

where R is the radius of the membrane. Again from $\int F\,ds$ we may set up the potential energy as

$$\frac{1}{2}\int_0^R\int_0^{2\pi}\left[T_1\left(\frac{\partial z}{\partial r}\right)^2+T_2\frac{1}{r^2}\left(\frac{\partial z}{\partial\theta}\right)^2\right]rdrd\theta$$

The integral converges at the lower limit for r as we shall see that these integrals are uniformly bounded in t.

Then the total energy E_0 of the system is equal to the sum of the kinetic and potential energies as

$$E_0=\frac{1}{2}\int_0^R\int_0^{2\pi}\left[T_1\left(\frac{\partial z}{\partial r}\right)^2+\frac{1}{r^2}T_2\left(\frac{\partial z}{\partial\theta}\right)^2+\mu\left(\frac{\partial z}{\partial t}\right)^2\right]rdrd\theta \qquad (2)$$

In rectangular coördinates this becomes

$$E_0=\frac{1}{2}\int_{x_0}^{x_1}\int_{y_0}^{y_1}\left[T_3\left(\frac{\partial z}{\partial x}\right)^2+T_4\left(\frac{\partial z}{\partial y}\right)^2+\mu_1\left(\frac{\partial z}{\partial t}\right)^2\right]rdrd\theta \qquad (2')$$

The total energy E_0 shall be assumed finite for the system.

We wish to show that the energy E_0 is independent of the time. In order to do this we shall compute $\frac{\partial E_0}{\partial t}$. But since the integrand is not regular at $r=0$ we shall consider $E_0^{(r_0)}$ which we define to be the expression for the energy E_0 but with lower limit r_0 instead of zero. Now the proper conditions are satisfied and we differentiate partially with respect to t under the integral sign and obtain after rewriting

$$\frac{\partial E_0^{(r_0)}}{\partial t}=\int_{r_0}^R\int_0^{2\pi}\left[T_1\left(\frac{\partial z}{\partial r}\right)\left(\frac{\partial^2 z}{\partial r\partial t}\right)+\frac{1}{r^2}T_2\left(\frac{\partial z}{\partial\theta}\right)\left(\frac{\partial^2 z}{\partial\theta\partial t}\right)+\mu\left(\frac{\partial z}{\partial t}\right)\left(\frac{\partial^2 z}{\partial t^2}\right)\right]rdrd\theta \qquad (3)$$

since $\frac{\partial\mu}{\partial t}=\frac{\partial T_1}{\partial t}=\frac{\partial T_2}{\partial t}=0$.

By substituting (1) in (3) we obtain

$$\frac{\partial E_0^{(r_0)}}{\partial t}=\int_{r_0}^R\int_0^{2\pi}\left[T_1\left(\frac{\partial z}{\partial r}\right)\left(\frac{\partial^2 z}{\partial r\partial t}\right)+\frac{1}{r^2}T_2\left(\frac{\partial z}{\partial\theta}\right)\left(\frac{\partial^2 z}{\partial\theta\partial t}\right)+\left(\frac{\partial z}{\partial t}\right)\left\{\frac{1}{r}\frac{\partial}{\partial r}\left(rT_1\frac{\partial z}{\partial r}\right)+\frac{1}{r^2}\frac{\partial}{\partial\theta}\left(T_2\frac{\partial z}{\partial\theta}\right)\right\}\right]rdrd\theta$$

$$= \int_{r_0}^{R}\int_{0}^{2\pi}\left[rT_1\frac{\partial z}{\partial r}\frac{\partial}{\partial r}\left(\frac{\partial z}{\partial t}\right)+\left(\frac{\partial z}{\partial t}\right)\left\{\frac{\partial}{\partial r}\left(rT_1\frac{\partial z}{\partial r}\right)\right\}\right]drd\theta$$

$$+\int_{r_0}^{R}\int_{0}^{2\pi}\left[\frac{1}{r}T_2\frac{\partial z}{\partial\theta}\frac{\partial}{\partial\theta}\left(\frac{\partial z}{\partial t}\right)+\left(\frac{\partial z}{\partial t}\right)\left\{\frac{\partial}{\partial\theta}\left(\frac{1}{r}T_2\frac{\partial z}{\partial\theta}\right)\right\}\right]drd\theta$$

$$= \int_{0}^{2\pi}\left\{\int_{r_0}^{R}\frac{\partial}{\partial r}\left[rT_1\frac{\partial z}{\partial r}\frac{\partial z}{\partial t}\right]dr\right\}d\theta$$

$$+\int_{r_0}^{R}\left\{\int_{0}^{2\pi}\frac{\partial}{\partial\theta}\left[\frac{1}{r}T_2\frac{\partial z}{\partial\theta}\frac{\partial z}{\partial t}\right]d\theta\right\}dr$$

$$= \int_{0}^{2\pi}\left\{rT_1\frac{\partial z}{\partial r}\frac{\partial z}{\partial t}\Big|_{r_0}^{R}\right\}d\theta+\int_{r_0}^{R}\left\{\frac{1}{r}T_2\frac{\partial z}{\partial\theta}\frac{\partial z}{\partial t}\Big|_{0}^{2\pi}\right\}dr$$

All of the factors in the second of the last two integrals are periodic in θ with period 2π. Thus the second integral vanishes identically. Again if $\frac{\partial z}{\partial r}$ is finite, which it is as we shall later see, for $r = 0$ and $r = R$, then we may let $r_0 \to 0$ in the first integral and the integrand will vanish at both limits since $\frac{\partial z}{\partial t} = 0$ for $r = 0$ and since there is a factor r.

Then $\frac{\partial E_0^{(r_0)}}{\partial t} \to 0$ as $r_0 \to 0$ and

$$\frac{\partial E_0}{\partial t} = 0 \tag{4}$$

which means E_0 is independent of t and therefore constant. This fact might easily have been predicted since there is no loss of energy from dissipation.

In rectangular coördinates this becomes

$$\frac{\partial E_0}{\partial t} = \int_{y_0}^{y_1}\left\{T_3\frac{\partial z}{\partial x}\frac{\partial z}{\partial t}\Big|_{x_0}^{x_1}\right\}dy+\int_{x_0}^{x_1}\left\{T_4\frac{\partial z}{\partial y}\frac{\partial z}{\partial t}\Big|_{y_0}^{y_1}\right\}dx$$

and both integrands are identically zero.

The kinetic and the potential energy are both non-negative since μ, T_1, and T_2 are positive. That is

$$\frac{1}{2}\int_{0}^{R}\int_{0}^{2\pi}\mu\left(\frac{\partial z}{\partial t}\right)^2 rdrd\theta \geqq 0$$

and

$$\frac{1}{2}\int_0^R\int_0^{2\pi}\left[T_1\left(\frac{\partial z}{\partial r}\right)^2+\frac{1}{r^2}T_2\left(\frac{\partial z}{\partial\theta}\right)^2\right]rdrd\theta \geqq 0\,.$$

Then from (2) both energies are bounded for all t. That is

$$\frac{1}{2}\int_0^R\int_0^{2\pi}\mu\left(\frac{\partial z}{\partial t}\right)^2 rdrd\theta \leqq E_0$$

and

$$\frac{1}{2}\int_0^R\int_0^{2\pi}\left[T_1\left(\frac{\partial z}{\partial r}\right)^2+\frac{1}{r^2}T_2\left(\frac{\partial z}{\partial\theta}\right)^2\right]rdrd\theta \leqq E_0\,.$$

Now if we take T as the minimum value of T_1 and T_2, and m the minimum value of μ, then certainly we have

$$\int_0^R\int_0^{2\pi}\left(\frac{\partial z}{\partial r}\right)^2 rdrd\theta+\int_0^R\int_0^{2\pi}\left(\frac{1}{r}\frac{\partial z}{\partial\theta}\right)^2 rdrd\theta \leqq \frac{2E_0}{T},$$

or, since both integrals are non-negative,

$$\int_0^R\int_0^{2\pi}\left(\frac{\partial z}{\partial r}\right)^2 rdrd\theta \leqq \frac{2E_0}{T}, \tag{5}$$

$$\int_0^R\int_0^{2\pi}\left(\frac{1}{r}\frac{\partial z}{\partial t}\right)^2 rdrd\theta \leqq \frac{2E_0}{T}, \tag{6}$$

and

$$\int_0^R\int_0^{2\pi}\left(\frac{\partial z}{\partial t}\right)^2 rdrd\theta \geqq \frac{2E_0}{m}. \tag{7}$$

If we differentiate (1) with respect to t we obtain after rewriting

$$\frac{1}{r}\frac{\partial}{\partial r}\left[rT_1\frac{\partial}{\partial r}\left(\frac{\partial z}{\partial t}\right)\right]+\frac{1}{r^2}\frac{\partial}{\partial\theta}\left[T_2\frac{\partial}{\partial\theta}\left(\frac{\partial z}{\partial t}\right)\right]=\mu\frac{\partial^2}{\partial t^2}\left(\frac{\partial z}{\partial t}\right) \tag{8}$$

which is (1) for $\frac{\partial z}{\partial t}$ instead of z. But $\frac{\partial z}{\partial t}$ is the velocity of the displacement, and for any physical membrane we may assume the higher energies $E_1, \cdots, E_n$ for the quantities $\frac{\partial z}{\partial t}, \cdots \frac{\partial^n z}{\partial t^n}$ to be existent and finite. Thus by differentiating (1) n times we obtain (1) for $\frac{\partial^n z}{\partial t^n}$ instead of z. Therefore we may set up the higher energies as

$$E_n=\frac{1}{2}\int_0^R\int_0^{2\pi}\left\{T_1\left(\frac{\partial z^{(n)}}{\partial r}\right)^2+\frac{1}{r^2}T_2\left(\frac{\partial z^{(n)}}{\partial\theta}\right)^2+\mu\left(\frac{\partial z^{(n)}}{\partial t}\right)^2\right\}rdrd\theta \tag{9}$$

where $z^{(n)} = \frac{\partial^n z}{\partial t^n}$. It is evident that we may show that $\frac{\partial E_n}{\partial t} = 0$ in exactly the same manner as for $\frac{\partial E_0}{\partial t} = 0$. From this we derive inequalities analogous to (5), (6), and (7) as

$$\int_0^R \int_0^{2\pi} \left(\frac{\partial z^{(n)}}{\partial r}\right)^2 r dr d\theta < A , \tag{10}$$

$$\int_0^R \int_0^{2\pi} \left(\frac{1}{r}\frac{\partial z^{(n)}}{\partial \theta}\right)^2 r dr d\theta < A , \tag{11}$$

and

$$\int_0^R \int_0^{2\pi} \left(\frac{\partial z^{(n)}}{\partial t}\right)^2 r dr d\theta < A \tag{12}$$

where A depends on E_n and on the minimum values of T_1, T_2, and μ. In rectangular coördinates (10), (11), and (12) become

$$\int\int_S \left(\frac{\partial z^{(n)}}{\partial x}\right)^2 dx dy < A \tag{13}$$

$$\int\int_S \left(\frac{\partial z^{(n)}}{\partial y}\right)^2 dx dy < A \tag{14}$$

and

$$\int\int_S \left(\frac{\partial z^{(n)}}{\partial t}\right)^2 dx dy < A \tag{15}$$

where the region of integration S is such that $|x| < R, |y| < R$, and where all functions are defined as vanishing identically outside the region for which $|r| < R$.

We will now find it convenient to continue the discussion in rectangular coördinates.

Any directional derivative $\frac{dz^{(n)}}{ds}$ satisfies the following inequality

$$\left(\frac{dz^{(n)}}{ds}\right)^2 \leqq 4\left[\left(\frac{\partial z^{(n)}}{\partial x}\right)^2 + \left(\frac{\partial z^{(n)}}{\partial y}\right)^2\right]$$

Therefore

$$\int\int_S \left(\frac{dz^{(n)}}{ds}\right)^2 dx dy \leqq 8A$$

and any directional derivative is bounded in the mean for all t, and this is true for the direction of maximum change.

If now we demand that T_3 be differentiable in x and T_4 in y (or in polar coördinates, T_1 in r and T_2 in θ), then we may expand (1′) to obtain

$$T_3\frac{\partial^2 z}{\partial x^2}+T_4\frac{\partial^2 z}{\partial y^2}+\frac{\partial T_3}{\partial x}\frac{\partial z}{\partial x}+\frac{\partial T_4}{\partial y}\frac{\partial z}{\partial y}=\mu_1\frac{\partial^2 z}{\partial t^2}$$

Then by transposing certain terms, squaring, multiplying by $dxdy$, integrating, and using the inequalities (13), (14), and (15) for $n = 0$ we obtain

$$\int\int_s\left(T_3\frac{\partial^2 z}{\partial x^2}+T_4\frac{\partial^2 z}{\partial y^2}\right)^2 dxdy$$

$$=\int\int_s\left[\mu_1\frac{\partial^2 z}{\partial t^2}-\frac{\partial T_3}{\partial x}\frac{\partial z}{\partial x}-\frac{\partial T_4}{\partial y}\frac{\partial z}{\partial y}\right]^2 dxdy\leq B'$$

where B' depends on E_0, E_1, and on the minimum values of T_3, T_4, and μ_1. Similarly

$$\int\int\left(T_3\frac{\partial^2 z^{(n)}}{\partial x^2}+T_4\frac{\partial^2 z^{(n)}}{\partial y^2}\right)^2 dxdy\leq B \tag{16}$$

where B depends on E_n, E_{n+1}, and on the minimum values of T_3, T_4, and μ_1.

It will be necessary to reduce the elliptic differential equation (1′) to its normal form as

$$\Delta z+a(x,y)\frac{\partial z}{\partial x}+b(x,y)\frac{\partial z}{\partial y}+c(x,y)z=g(x,y)\frac{\partial^2 z}{\partial t^2} \tag{17}$$

where z is the displacement as a function of time t and space coördinates x and y in terms of the new coördinates, and where Δz is the Laplacian. From the properties of T_3, T_4, and μ_1, the coefficients a, b, c, and g are absolutely bounded though not necessarily non-negative. The region of integration S becomes S'.

Now for any $t = t_1$ set

$$L_1(z)\equiv\Delta z=-2\pi\sigma_1(t_1;x,y) \tag{18}$$

According to the theory of elliptic linear partial differential equations

of second order[9] we have

$$z(t_1; x, y) = \int\int_{S'} G_1(x, y; \xi, \eta)\, \sigma_1(t_1; \xi, \eta)\, d\xi d\eta + v_1(x, y)$$

where t_1 is taken as a parameter, where G_1 is the Green's function belonging to $L_1(z)$ and does not depend on t_1 since T_3 and T_4 do not, and where v_1 is a bounded regular potential function which takes on certain values on the boundary of S'. From the Schwarz inequality

$$z(t_1; x, y) \leqq \left[\int\int_{S'} G_1^2\, d\xi d\eta\right]^{\frac{1}{2}} \cdot \left[\int\int_{S'} \sigma_1^2\, d\xi d\eta\right]^{\frac{1}{2}} + v\,.$$

The integral of the square of the Green's function is finite and the second integral is bounded by (16). Since (1′) holds as well for all $z^{(n)}(t; x, y)$, this argument may be repeated to show that all $z^{(n)}$ are bounded in x, y and t. That is

$$z^{(n)}(t; x, y) < D, \qquad n = 0, 1, 2, \cdots, n \tag{19}$$

Now again for any $t = t_1$ set up (17) as

$$L_2(z) \equiv \Delta z + a\frac{\partial z}{\partial x} + b\frac{\partial z}{\partial y} + cz = g\frac{\partial^2 z}{\partial t^2} = -2\pi\sigma_2(t_1; x, y) \tag{20}$$

Again using the Green's function for $L_2(z)$ we have from the theory

$$z(t_1; x, y) = \frac{1}{2\pi}\int\int_{S'} \Big\{G_2(x, y; \xi, \eta)\, c(\xi, \eta) - \frac{\partial}{\partial \xi}\left[a(\xi, \eta)\, G_2(x, y; \xi, \eta)\right] - \frac{\partial}{\partial \eta}\left[b(\xi, \eta)\, G_2(x, y; \xi, \eta)\right]\Big\} z(t_1; \xi, \eta)\, d\xi d\eta + \int\int_{S'} G_2(x, y; \xi, \eta)\, \sigma_2(t_1; \xi, \eta)\, d\xi d\eta + v_2(x, y) \tag{21}$$

where G_2 is the Green's function for $L_2(z)$ and v_2 is again a bounded regular potential functiom. From potential theory we know that the Green's function here has a logarithmic singularity at the pole x, y, and has the form

$$\operatorname{Log}\frac{1}{r} + w(x, y; \xi, \eta)$$

where $r = [(x - \xi)^2 + (y - \eta)^2]^{\frac{1}{2}}$, and where w is harmonic in S'.

[9] See bibliography 7.

From the symmetry of the Green's function

$$\frac{\partial G_2}{\partial \xi} = -\frac{\partial G_2}{\partial x} \quad \text{and} \quad \frac{\partial G_2}{\partial \eta} = -\frac{\partial G_2}{\partial y}.$$

Now we may obtain $\frac{\partial z}{\partial x}$ by differentiating (21) under the integral sign with respect to x, for from the symmetry of the Green's function we may integrate by parts each integral of the x derivative of (21). The first part of this integration by parts gives a difference at the end values of ξ which are different from the pole at x, y, so that the integral remains finite. The second part of the integration gives integrals like those of (21), but involving derivatives of a, b, and c instead of these coefficients themselves. Then if these derivatives of a, b, and c exist and are bounded $\frac{\partial z}{\partial x}$ exists, is bounded, and is properly obtained by differentiating (21) under the integral sign. The existence of the first derivatives of a, b, and c is equivalent to the existence of the second derivatives of T_3 and T_4. Similarly, we may obtain $\frac{\partial^2 z}{\partial x^2}$ and higher derivatives with respect to x by differentiating under the integral sign, if only we demand bounded second and higher derivatives of a, b, and c. Then in a like manner we may show the boundedness of $\frac{\partial z}{\partial y}$, $\frac{\partial^2 z}{\partial y^2}$, and higher derivatives, and also cross derivatives, by the same restrictions on the coefficients a, b, and c and, consequently, on T_3 and T_4. For the purpose of our problem as presented here we need only up to and including the second derivatives of z with respect to x and y. Still higher derivatives would be required if we were to consider almost periodic properties of the curvatures of the membrane. However, we shall consider almost periodicity only in the displacement, the slope, and the time derivatives of the displacement.

LEMMA. If $f(t; x, y)$ is developable in a Fourier series and if

$$\int\int_s \left| \frac{\partial f(t; x, y)}{\partial x} \right|^2 dxdy$$

is bounded for all t, then the remainder of the series is less than a quantity depending on $\left(\frac{\partial f}{\partial x}\right)^2$ and on the number of terms taken in the series, and vanishes uniformly as the number of terms becomes infinite.

From the proof it is evident that the lemma is true also for the y derivative.

Let $f(t; x, y)$ be expanded in the series

$$f(t; x, y) \sim R\left\{\sum_{n,m=-\infty}^{\infty} a_{nm}(t)\, e^{i\frac{\pi}{R}(nx+my)}\right\}$$

where

$$a_{nm}(t) = \frac{1}{4R^2}\int_{-R}^{R}\int_{-R}^{R} f(t; x, y)\, e^{-i\frac{\pi}{R}(nx+my)}\, dxdy$$

and where t is considered as a parameter. Upon differentiating we obtain

$$\frac{\partial f}{\partial x} = R\left\{\sum_{n,m=-\infty}^{\infty} a_{nm}(t)\,\frac{n\pi}{R}\, ie^{i\frac{\pi}{R}(nx+my)}\right\}$$

Then because of orthogonality

$$\iint_s \left(\frac{\partial f}{\partial x}\right)^2 dxdy = \iint_s \left(R\sum_{n,m} a_{nm}(t)\,\frac{n\pi}{R}\, ie^{i\frac{\pi}{R}(nx+my)}\right)^2 dxdy$$

$$= \iint_s \left[\sum_{n,m} (Ra_{nm})^2\,\frac{\pi^2}{R^2}\, n^2 \sin^2\frac{\pi}{R}(nx+my)\right] dxdy$$

$$+ \iint_s \left[\sum_{n,m} \left(Ia_{nm}\right)^2\,\frac{\pi^2}{R^2}\, n^2 \cos^2\frac{\pi}{R}(nx+my)\right] dxdy$$

$$= \frac{\pi^2}{R^2}\sum_{n,m} n^2\left\{(Ra_{nm})^2\iint_s \sin^2\frac{\pi}{R}(nx+my)\, dxdy\right.$$

$$\left. + (Ia_{nm})^2\iint_s \cos^2\frac{\pi}{R}(nx+my)\, dxdy\right\}$$

$$= \frac{\pi^2}{R^2}\sum_{n,m} n^2\left\{(Ra_{nm})^2 2R^2 + (Ia_{nm})^2 2R^2\right\} = 2\pi^2\sum_{n,m=-\infty}^{\infty} n^2 \mid a_{nm}\mid^2$$

where the integration over S is over the square of side $2R$. Since by hypothesis $\dfrac{\partial f}{\partial x}$ is bounded in the mean we may write

$$\sum_{n,m=-\infty}^{\infty} \mid a_{nm}\mid^2 \cdot n^2 < A$$

where A depends only on the bound of the average value of $\left(\frac{\partial f}{\partial x}\right)^2$. It follows immediately that

$$\sum_{N,M}^{\infty} |a_{nm}|^2 < \frac{A}{N^2}$$

and

$$\sum_{-\infty}^{-N,-M} |a_{nm}|^2 < \frac{A}{N^2}$$

Consequently

$$\sum_{-N,-M}^{N,M} |a_{nm}|^2$$

is uniformly bounded in t, N, and M, and the remainder of the series is less that $2A/N^2$ which vanishes as the number of terms taken becomes infinite in both m and n, and the lemma is proved.

We are now ready to prove that in a finite time the system in the mean returns to within ϵ of its original position.

Define the distance between two points t_1 and t_2 by

$$D_{12} = \int\int_s \left[|z_2 - z_1|^2 + \left|\frac{dz_2}{ds} - \frac{dz_1}{ds}\right|^2 + \left|\frac{\partial z_2}{\partial t} - \frac{\partial z_1}{\partial t}\right|^2 + \left|\frac{\partial^2 z_2}{\partial t^2} - \frac{\partial^2 z_1}{\partial t^2}\right|^2 \right] dxdy$$

Here we mean by distance the sum of the differences in the mean between displacement, slope, velocity, and acceleration of the vibrating system for any two values of time t_1 and t_2. Evidently we may include in our distance also higher time derivatives of the displacement to any order where energies to that order have been assumed to exist and be finite.

For convenience we shall make a slight modification in the definition for distance. If we set

$$\frac{dz}{ds} = \frac{\partial z}{\partial x}\frac{dx}{ds} + \frac{\partial z}{\partial y}\frac{dy}{ds}$$

then

$$\left|\frac{dz_2}{ds} - \frac{dz_1}{ds}\right|^2 = \left|\frac{\partial z_2}{\partial x}\frac{dx}{ds} + \frac{\partial z_2}{\partial y}\frac{dy}{ds} - \frac{\partial z_1}{\partial x}\frac{dx}{ds} - \frac{\partial z_1}{\partial y}\frac{dy}{ds}\right|^2$$

$$\leqq 2\left|\frac{dx}{ds}\right|^2\cdot\left|\frac{\partial z_2}{\partial x}-\frac{\partial z_1}{\partial x}\right|^2+2\left|\frac{dy}{ds}\right|^2\cdot\left|\frac{\partial z_2}{\partial y}-\frac{\partial z_1}{\partial y}\right|^2$$

$$\leqq 2\left|\frac{\partial z_2}{\partial x}-\frac{\partial z_1}{\partial x}\right|^2+2\left|\frac{\partial z_2}{\partial y}-\frac{\partial z_1}{\partial y}\right|^2.$$

Therefore

$$D_{12}\leqq D'_{12}=\int\int_s\left[\,|z_2-z_1|^2+2\left|\frac{\partial z_2}{\partial x}-\frac{\partial z_1}{\partial x}\right|^2+2\left|\frac{\partial z_2}{\partial y}-\frac{\partial z_1}{\partial y}\right|^2\right.$$

$$\left.+\left|\frac{\partial z_2}{\partial t}-\frac{\partial z_1}{\partial t}\right|^2+\left|\frac{\partial^2 z_2}{\partial t^2}-\frac{\partial^2 z_1}{\partial t^2}\right|^2\right]dxdy.$$

It follows from (13), (14), and (15) that the distance for any two choices of time is uniformly bounded.

We know from the lemma that

$$\sum_{-N,-M}^{N,M}|a_{mn}(t)|^2$$

is uniformly bounded in t, M, and N in the series for z, the displacement. We may expand the other four elements of the distance D'_{12} in similar series and by using the lemma show that their respective series are uniformly bounded in t, M, and N. Then all the numbers $|a_{mn}(t)|$ are uniformly bounded in t.

Take $M=N$. We define a hyperspace of $5(2N+1)^2$ dimensions. For each t the $5(2N+1)^2$ values of $|a_{mn}|$ determined define a point in this hyperspace where distance is defined according to D'_{12}. All of the points so determined lie in a finite part of this hyperspace since D'_{12} is bounded for all t_1 and t_2. Now let this finite part be divided into compartments, say as hypercubes, of maximum dimension $\leqq\epsilon$, where ϵ is an arbitrarily chosen small positive number. For a fixed ϵ the number of such compartments is finite. Furthermore, with each compartment there is associated a certain combination of the elements z, $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, $\frac{\partial z}{\partial t}$ and $\frac{\partial^2 z}{\partial t^2}$. Then each compartment represents a state of the vibrating system. To each compartment attach a number which tells the first value of time for which the vibrating system corresponds to that compartment; that is, the time for which the values of the five elements of the distance have simultaneously come close to the values of the respective elements as represented by the particular compartment. Now since there are a finite number of compartments we may take the largest

time of all these and by that time the system will have entered, that is, corresponded to, every compartment that it ever possibly may (where those compartments not reached in finite time are said never to be reached at all). Thus, if we have given an $\epsilon > 0$, then there always exists a finite time before which the system will have been closer than ϵ to any position that ever it may possible approach in displacement, slope, velocity, and acceleration, simultaneously.

Now from the definition of a function almost periodic in the mean this largest value of time as above may be interpreted as the initial inclusion interval, depending on ϵ. That is, for any given time t_2 there is always a value of time $t_1 < l$ such that at t_1 the distance according to our definition is within ϵ in the mean of the system at the given time t_2.

If now for any element, say the displacement, the values for two different times t_1 and t_2 differ in the mean by a small quantity ϵ, then they will so remain during successive vibrations of the membrane. That is, if

$$\int\int_s [z(t_2; x, y) - z(t_1; x, y)]^2 \, dxdy \leqq \epsilon,$$

then

$$\int\int_s [z(t_2 + t; x, y) - z(t_1 + t; x, y)]^2 \, dxdy \leqq k\epsilon$$

uniformly for all t where k is a constant independent of t. This is so, since from the discussion if the displacements are close for t_1 and t_2 then so are all the time derivatives to any given order, and therefore the displacements remain close at any later times $t_1 + t$ and $t_2 + t$.

Thus we have shown the *existence* of *displacement numbers*, since by definition $t_2 - t_1$ is a displacement number.

Now since for any time t_2 in any interval from T to $T + l$ there is a time $t_1 < l$ such that the positions of the system at t_2 and at t_1 differ in the mean by a small quantity, we may write

$$\int\int_s [z(T + \theta_1 l; x, y) - z(\theta_2 l; x, y)]^2 \, dxdy \leqq \epsilon$$

for all T wher $0 < \theta_1, \theta_2 < 1$. Then $T + \theta_1 l - \theta_2 l$ is a displacement number, and since T is arbitrary these exist in an infinite number with a *relatively dense* distribution having an inclusion interval l.

Therefore we have proved that the *displacement* is *almost periodic in the mean in t*. But this argument holds as well for the slope, the veloc-

ity, the acceleration, and any higher time derivative of the displacement, consequently these elements too are almost periodic in the mean in t, all with the same inclusion interval l, depending on a given ϵ.

III. Conclusion

In completing this paper we shall consider the characteristic solutions, or *eigenfunktionen* as they are often called. To do this we shall study the integral equation associated with the differential equation of the membrane (1′) in rectangular coördinates. In polar coördinates the discussion is similar.

By expanding the differential equation (1′) and integrating twice each with respect to x, y, and t, using direct integration and integration by parts, we obtain the following integral equation:

$$\int_0^t\int_0^t\int_{v_0}^{v}\int_{v_0}^{v}\Big[T_3(x,y)\,z(t;x,y) - T_3(x_0,y)\,z(t;x_0,y) - T_3(x_0,y)\left(\frac{\partial z}{\partial x}\right)_{x_0}(x-x_0)\Big]\,dy^2dt^2$$

$$+\int_0^t\int_0^t\int_{x_0}^{x}\int_{x_0}^{x}\Big[T_4(x,y)\,z(t;x,y) - T_4(x,y_0)\,z(t;x,y_0) - T_4(x,y_0)\left(\frac{\partial z}{\partial y}\right)_{y_0}(y-y_0)\Big]\,dx^2dt^2$$

$$-\int_0^t\int_0^t\int_{v_0}^{v}\int_{v_0}^{v}\int_{x_0}^{x}\frac{\partial T_3}{\partial x}\,z(t;x,y)\,dt^2dy^2dx - \int_0^t\int_0^t\int_{x_0}^{x}\int_{x_0}^{x}\int_{v_0}^{v}\frac{\partial T_4}{\partial y}\,z(t;x,y)\,dt^2dx^2dy$$

$$=\int_{x_0}^{x}\int_{x_0}^{x}\int_{v_0}^{v}\int_{v_0}^{v}\mu(x,y)\left[z(t;x,y) - z(t_0;x,y) - \left(\frac{\partial z}{\partial t}\right)_{t_0}(t-t_0)\right]\,dx^2dy^2.$$

Any solution of (1′) is a solution of (23). Therefore, the approximating trigonometric polynomial of the almost periodic in the mean solution of the differential equation (1′) is also a solution to (23). Furthermore, the uniform limit of a sequence of solutions to an integral equation is also a solution. We now wish to show that each separate term of the approximating trigonometric polynomial is a solution to the integral equation (23).

If $f(t;x,y)$ is a solution to (23), then $f(t+t_0;x,y)$ is also a solution

where t_0 is any time. This is so since it represents only a displacement in the time and time does not appear explicitly in the equation. Since a constant times a solution is also a solution, then $k(t_0) \cdot f(t + t_0; x, y)$ is a solution. Furthermore, from properties of functions almost periodic in the mean and from the existence of the mean value

$$\int_{-A}^{A} k(t_0)\, f(t + t_0; x, y)\, dt_0$$

and

$$\text{l.i.m.}_{A \to \infty} \frac{1}{2A} \int_{-A}^{A} k(t_0)\, f(t + t_0; x, y)\, dt_0 \tag{24}$$

are also solutions, where $k(t_0)$ must be taken in absolute value less than or equal to 1 so that the existence of the mean value would imply the existence of (24).

Take $k(t_0) = e^{i\Lambda t_0}$ where Λ is any of the Fourier exponents of $f(t; x, y)$. Then from (24) $M\{e^{i\Lambda t_0} f(t + t_0; x, y)\}$ is a solution where M is the mean value in t_0.

However, this mean value is really only one term when $f(t + t_0; x, y)$ is taken in its trigonometric series, since the averaging process causes every term to vanish except that term in the series for $f(t + t_0; x, y)$ with exponent Λ. Then if Λ is taken as Λ_k of the expansion

$$M\{e^{i\Lambda_k t_0} f(t + t_0; x, y)\} = A_k(x, y)\, e^{i\Lambda_k t}\,.$$

But this is a single term of the approximating polynomial of $f(t; x, y)$ and we shall show that these single terms satisfy the differential equation. If we substitute $z = A_k(x, y)\, e^{i\Lambda_k t}$ in (23) and differentiate twice each with respect to t, x, and y we obtain the differential equation (1′). The only question is that of the existence of the necessary derivatives. All the necessary derivatives in t exist since we may differentiate $e^{i\Lambda_k t}$ as often as we please. We must investigate $A_k(x, y)$. First differentiate (23) twice with respect to t. Then differentiate once with respect to y. This will give only one derivative $\frac{\partial A_k}{\partial y}$; everything else in the equation exists, therefore, $\frac{\partial A_k}{\partial y}$ exists. Differentiate once again with respect to y to get only one term in $\frac{\partial^2 A_k}{\partial y^2}$ and as before this must exist since everything else exists. Similarly we may show the existence first of $\frac{\partial A_k}{\partial x}$ then of $\frac{\partial^2 A_k}{\partial x^2}$.

Thus single terms of the approximating trigonometric polynomial solutions are themselves solutions.

Then a solution of the differential equation (1′) may be uniformly approximated by a series as

$$\sum_{k=1}^{N} A_k(x, y)\, e^{i\Lambda_k t}$$

reducing for $t = 0$ to the initial position $\varphi(x, y)$, and where $A_k(x, y)\, e^{i\Lambda_k t}$ separately satisfies (1′). Likewise, the solution to (1) may be approximated, reducing for $t = 0$ to the initial position $\varphi_1(r, \theta)$.

These $A_k(x, y)$ are the characteristic functions, or eigenfunktionen. They satisfy the differential equation

$$\frac{d}{dx}\left(T_3 \frac{\partial z}{\partial x}\right) + \frac{d}{dy}\left(T_4 \frac{\partial z}{\partial y}\right) = -\mu\Lambda_k^2 z$$

The initial position of the membrane must be given closely by

$$\sum_{k=1}^{N} A_k(x, y)$$

That is, we may take

$$\int\int_S \left| \varphi(x, y) - \sum_{k=1}^{N} A_k(x, y) \right|^2 dxdy$$

uniformly as small as we please by properly choosing N each time.

These characteristic functions represent the standing waves of the membrane and they give the position of the nodal curves.

BIBLIOGRAPHY

1. A. Besicovitch. "Almost Periodic Functions." Cambridge, 1932 (Bibliography).
2. A. Besicovitch and H. Bohr. "Almost Periodicity and General Trigonometric Series." Acta. Math., vol. 57, pp. 203–292 (1931).
3. A. Besicovitch. "Analysis of Conditions of Generalized Almost Periodicity." Acta. Math., vol. 58, pp. 217–230.
4. S. Bochner. "Fastperiodische Lösungen der Wellengleichung." Acta Math., vol. 62, pp. 227–237 (1933).
5. H. Bohr. "Zur Theorie der Fastperiodische Funktionen." Acta Math., vol. 45, pp. 29–127 (1924).
   ——— "Zur Theorie der Fastperiodische Funktionen." Acta Math., vol. 46–47 (1925–1926).
6. H. Bohr. "Fastperiodische Funktionen." Ergebnisse der Mathematik, band 1 (5) (1932).

7. L. Lichtenstein. Encyklopädie der Mathematische Wissenschaften, vol. 2, 2nd part, 2nd half, p. 1282.
8. C. Muckenhoupt. "Almost Periodic Functions and Vibrating Systems." Journal of Math. and Physics, M. I. T., vol. VIII, no. 3 (1929) (Bibliography).
9. H. Weyl. "Integralgleichungen und Fastperiodische Funktionen." Math. Ann., band 97, pp. 338–356 (1926).
10. S. Bochner. "Beitrage zur Theorie der fastperiodische Funktionen." Math. Ann., band 96, pp. 119–147 (1926).

For a more complete list see bibliography at end of paper by N. Wiener; "Generalized Harmonic Analysis," Acta Math., vol. 55 (1930). See also (1) and (8) above.